

ISKANJE OBRAZOV NA OSNOVI BARV S POMOČJO STATISTIČNIH METOD RAZPOZNAVANJA VZORCEV

Borut Batagelj

MAGISTRSKA NALOGA

predložena

Fakulteti za računalništvo in informatiko

Univerze v Ljubljani

kot delna izpolnitev pogoja za pridobitev naslova
magister računalništva in informatike

Ljubljana, 2004

Mentor:

prof. dr. Franc Solina

Magistrska naloga je bila izdelana pod mentorstvom prof. dr. Franca Soline in je last Fakultete za računalništvo in informatiko v Ljubljani. Za objavlanje in uporabo rezultatov magistrskega dela je potrebno soglasje zgoraj omenjene ustanove.

Besedilo je oblikovano s sistemom za pripravo besedil \LaTeX .

Povzetek

V zadnjem času postaja video nadzor vse pomembnejši in s tem tudi sistemi za iskanje in prepoznavo človeških obrazov na slikah. Zato se v magistrskem delu ukvarjam s problemom iskanja obrazov na slikah.

Pri metodah za iskanje obrazov na podlagi barve smo velikokrat omejeni na človeške obraze samo določene polti, same metode pa so tudi zelo odvisne od osvetlitve. V magistrskem delu zato poskušam s pomočjo kromatičnega barvnega prostora odvisnost od osvetlitve zmanjšati. Preizkusil bom različne metode za barvno segmentacijo na osnovi parametričnega in neparametričnega modela. S pomočjo teh modelov bom poskušal modelirati kožno barvo pri različnih osvetlitvah in različnih kožnih polteh. Uspešnost metod bom primerjal z metodo, ki deluje v barvnem prostoru RGB na osnovi eksplicitno določenih mej.

Za potrjevanje označenih kožnih regij bom uporabil metodo na osnovi videza, ki nam med vsemi metodami obljublja najboljše rezultate. Izdelal in preizkusil bom metodo BDF, ki na osnovi naučenega vzorca obraza in neobraza s pomočjo Bayesovega klasifikatorja najde frontalne obraze na sivinskih slikah.

Glavna slabost metod na osnovi videza je njihova časovna zahtevnost, zato bom poskušal izdelati metodo, ki bo kombinirala pristop na osnovi barv in pristop na osnovi videza. S pomočjo tako izdelane metode bom poskušal doseči hitro in učinkovito iskanje frontalnih obrazov na barvnih slikah.

Ključne besede: iskanje človeških obrazov, barvna segmentacija, kožna barva, barvni prostori, razpoznavanje vzorcev, statistična metoda, metoda na osnovi barv, metoda na osnovi videza, metoda BDF

Abstract

In recent years, the importance of video surveillance has been growing. The same is true of systems for human face detection and recognition. Therefore, the following Master's degree deals with the human face detection in color images.

When using the color-based face detection methods one is repeatedly limited to human faces of certain complexions only and the methods themselves are often dependent on the lighting conditions. Therefore, I tried to diminish the dependence on the lighting conditions by means of chrominance color space. Various methods for color segmentation are going to be tested on the basis of parametric and non-parametric model. By means of these models the skin color is going to be modelled at different exposures and skin complexions. The efficiency of these methods will be compared to the one which functions in RGB color space on the basis of explicitly defined skin regions.

To confirm the selected skin regions, the appearance-based method, which provides us with the best results, will be used. The Bayesian Discriminating Features (BDF) method for multiple frontal face detection by integrating feature analysis, modeling, and the Bayes classifier will be introduced and tested.

The main disadvantage of the appearance-based methods is their temporal demand, therefore, the method which will combine the color- and appearance-based approach, will be used. By means of such a method, I will try to achieve fast and efficient face detection in color images.

Key words: human face detection, color segmentation, skin color, color spaces, pattern recognition, statistical method, color-based method, appearance-based method, BDF method

Kazalo

Kazalo slik	xviii
Kazalo tabel	xix
1 Uvod	1
2 Pregled različnih tehnik za iskanje obrazov na sliki	9
2.1 Pristop na osnovi znanja - od zgoraj navzdol	11
2.2 Pristop na osnovi značilnk - od spodaj navzgor	13
2.2.1 Metode na osnovi značilnk obrazov	13
2.2.2 Metode na osnovi teksture obraza	16
2.2.3 Metode na osnovi barve kože	17
2.2.4 Metode na osnovi kombinacije različnih značilnk	18
2.3 Pristop na osnovi šablone	19
2.3.1 Pred-pripravljene šablone	20
2.3.2 Spremenljive šablone	20
2.4 Pristop na osnovi videza	21
2.4.1 Lastni obrazi (ang. Eigenfaces)	22
2.4.2 Metode na osnovi porazdelitev	23
2.4.3 Nevronske mreže	25
2.4.4 Metoda podpornih vektorjev	26
2.4.5 Naivni Bayesov klasifikator	27
2.5 Zaključek	28

3	Različne metode in barvni prostori za segmentacijo kožne barve	29
3.1	Pregled barvnih prostorov	29
3.1.1	Barvni prostor RGB (ang. Red, Green, Blue)	30
3.1.2	Normaliziran barvni prostor RGB ali barvni prostor rg	31
3.1.3	HSI, HSV, HSL - Hue, Saturation, Intensity (Value, Lightness) . . .	32
3.1.4	TSL - Tint, Saturation, Lightness	35
3.1.5	Y'CrCb barvni prostor	36
3.1.6	Barvna prostora CIE-Luv in CIE-Lab	37
3.1.7	Razmerja med RGB kanali	41
3.2	Modeliranje kožne barve	41
3.2.1	Eksplisitno določene meje	41
3.2.2	Neparametrično modeliranje	42
3.2.3	Parametrične metode	45
3.3	Izbira prave metode in ustreznega barvnega prostora	49
3.3.1	Povzetek različnih metod	49
3.3.2	Povzetek različnih barvnih prostorov	50
3.3.3	Izbira primerne barvnega prostora	51
3.3.4	Preizkus različnih metod	52
4	Implementacija različnih metod v izbranem barvnem prostoru	53
4.1	Metoda eksplisitno določenih mej	53
4.2	Metoda na osnovi parametričnega modela	55
4.2.1	Postopek preizkusa uspešnosti metode	60
4.2.2	Modeliranje kožne barve pod različnimi osvetlitvami	62
4.2.3	Modeliranje kožne barve različnih kožnih polti	65
4.3	Metoda na osnovi neparametričnega modela	67
4.4	Zaključek	68
5	Iskanje obrazov na podlagi vzorcev	75
5.1	Bayesova metoda diskriminantnih značilk	76
5.2	Normalizacija učnih slik	77
5.3	Analiza diskriminantnih značilk	80

5.4	Statistična modela obrazov in neobrazov	82
5.4.1	Model obraza	82
5.4.2	Model neobraza	84
5.5	Bayesov klasifikator za iskanje frontalnih obrazov	85
5.6	Preizkus metode	87
5.6.1	Statistično učenje metode BDF	88
5.6.2	Rezultati preizkusa	89
5.6.3	Primerjava uspešnosti metode	91
5.7	Zaključek	91
6	Združitev metode na osnovi barve in metode BDF	101
6.1	Postopek iskanja obraza	102
6.2	Preizkus metode in rezultati	105
7	Zaključek in nadaljnje delo	111
7.1	Nadaljnje delo	113
	Zahvala	115
	Izjava	127

Slike

1.1	Različni primeri slike obrazov: a) obrazi v različnih pozah, b) različnih velikosti, c) z različnimi izrazi na obrazu, d) različno osvetljeni in e) obrazi različnih polti (vir: podatkovna baza FERET [49] in podatkovna baza Yale [20]).	2
1.2	Merila pri iskanju obrazov: a) vhodna slika, b) izbrani kandidati za obraz, c) najdeni obrazi: pravilno najden obraz, napačno najden obraz in obraz, ki ni bil najden.	6
2.1	Primer obraza, ki se uporablja pri pristopu na osnovi znanja: pravila so definirana na osnovi intenzivnosti slikovnih elementov obraznih značilnk. . .	12
2.2	Postopek klasifikacije, ki jo je uporabil Lenitis s sodelavci [36]. Testna slika se klasificira na podlagi modela oblike in izračunanih parametrov oblike ter parametrov intenzivnosti slikovnih točk.	21
2.3	Porazdelitvena modela za razred obrazov in neobrazov, ki sta ju uporabila Sung in Poggio v svojem delu [66]. Posamezen model sestavlja mešanica Gaussovih funkcij. Na desni strani so prikazani centri Gaussovih funkcij. .	24
2.4	Diagram nevronske mreže, ki jo je uporabil Rowley [53] v svoji metodi za iskanje obrazov na sliki.	26
3.1	Predstavitev barvnega prostora RGB s pomočjo kocke.	30
3.2	Predstavitev barvnih prostorov HSL in HSV.	34
3.3	CIE kromatični diagram za prikaz barv v Yxy barvnem prostoru. Trikotnik označuje barvni prostor RGB.	39

4.1	Nepravilno ohranjena barva.	55
4.2	Kožna barva, ki zaradi nestandardne osvetlitve ni bila prepoznana.	55
4.3	Primer obraza in označene kožne regije.	56
4.4	Porazdelitev kožne barve v kromatičnem barvnem prostoru.	56
4.5	Porazdeljenost kožne barve in pripadajoč Gaussov model.	57
4.6	Grafa funkcij TP in TN za določitev mejne vrednosti Φ_p	59
4.7	Določitev kožne barve na osnovi barvnega modela: a) vhodna slika, b) binarna slika z ohranjenimi področji kožne barve.	60
4.8	Primer slike z ročno označeni obrazi.	61
4.9	Postopek preverjanja uspešnosti metode.	62
4.10	Primer slike iz različnih podatkovnih baz.	63
4.11	Primer istega obraza pod različnimi osvetlitvami.	63
4.12	Porazdelitev kožne barve in pripadajoči Gaussov model za različne osvetlitve posamezne testne baze v kromatičnem barvnem prostoru. Posamezni pari slik pripadajo posamezni podatkovni bazi. Slika levo zgoraj in slika pod njo prikazujeta porazdelitev in Gaussov model za slike prve baze. . . .	70
4.13	Primer ohranjanja kožne barve na testni sliki nestandardne osvetlitve s pomočjo dveh metod: a) metoda eksplicitno določenih mej, b) metoda na osnovi parametričnega modela.	71
4.14	Obrisi Gaussovih krivulj modelov kožne barve slik različnih podatkovnih baz.	71
4.15	Primeri obrazov različnih rasnih skupin (vir: podatkovna baza FERET [49]).	72
4.16	Izložene barve obrazov različnih rasnih skupin.	72
4.17	Kožna barva predstavljena z dvodimenzionalnim histogramom.	72
4.18	Primer slike segmentirane s pomočjo metod: a) eksplicitno določenih mej, b) na osnovi neparametričnega modela, c) na osnovi neparametričnega modela, kateremu smo dodali model neuspešno klasificirane rdeče barve tal. . . .	73
5.1	Primeri obrazov iz podatkovne baze FERET.	78
5.2	Primer slike narave, ki ne vsebuje obrazov.	78
5.3	Učni primeri istih oseb iz množice obrazov.	79

5.4	Postopek normalizacije slike obraza: a) originalna slika obraza z označenimi koordinatami oči in nagibom glave, b) ustrezno zasukan obraz, c) zmanjšana slika obraza, d) izrezan osrednji del obraza glede na pozicijo oči.	79
5.5	Primeri normaliziranih obrazov učne množice.	80
5.6	Diskriminantna analiza značiln povprečnega obraza in neobraza.	81
5.7	Primeri najdenih obrazov iz testne baze SET1.	93
5.8	Primeri najdenih obrazov iz testne baze SET2.	94
5.9	Primer iskanja različno velikih obrazov.	95
5.10	Primer iskanja različno nagnjenih obrazov.	96
5.11	Primer uspešno najdenega velikega obraza.	97
5.12	Primer uspešno najdenega majhnega obraza.	97
5.13	Primer iskanja obrazov na sliki slabe kvalitete.	98
5.14	Primer delno zakritega obraza.	98
5.15	Izbrani učni primeri neobrazov na sliki iz narave.	99
5.16	Primer neuspešno najdenega obraza ter napačno najdenega obraza.	99
6.1	Postopek ohranjanja kožne barve: a) pomanjšana vhodna slika, b) ohranjena kožna barva združena v regije, c) ohranjene regije, ki zadoščajo pravilom.	103
6.2	Vhodna slika z označenimi ohranjenimi kožnimi regijami.	104
6.3	Maske ohranjenih regij, ki predstavljajo ohranjeno kožno barvo.	104
6.4	Ohranjena regija: a) pri različnih nagibih in b) različnih velikostih.	105
6.5	Primer dveh različno velikih in različno nagnjenih obrazov, ki jih je metoda uspešno našla.	106
6.6	Primer slike z uspešno najdenima obrazoma.	107
6.7	Primer slike z ročno označenimi frontalnimi obrazi.	108
6.8	Primer slik z uspešno najdenimi obrazi: a) različno nagnjena glava, b) delno obrnjena glava proti profilnemu položaju, c) obrazi različnih velikosti, d) delno zakrit obraz.	108
6.9	Primer slik, na katerih metoda ni našla obraza: a) obraz je premajhen, b) obraz je preveč obrnjen, c) prevelik odsev očal in d) kožna regija je napačno označena.	109
6.10	Primer slike, kjer je metoda nepravilno našla obraz.	109

6.11 Primer prevelike ohranjene kožne regije glede na velikost obraza.	110
--	-----

Tabele

1.1	Merila, ki se uporabljajo pri sistemih za iskanje obrazov na sliki.	5
2.1	Preizkus različnih metod za iskanje obrazov na testni množici MIT in testni množici CMU.	28
4.1	Lastnosti posameznih podatkovnih baz barvnih slik obrazov.	64
4.2	Preizkus metode na osnovi različnih parametričnih modelov na slikah različnih osvetlitev.	64
4.3	Preizkus metode na osnovi posplošenega parametričnega modela nad slikami različnih osvetlitev.	66
4.4	Preizkus metode na osnovi različnih neparametričnih modelov na slikah različnih osvetlitev.	67
5.1	Testne baze: SET1, SET2 in SET3.	88
5.2	Rezultati preizkusa na testnih bazah: SET1, SET2 in SET3.	90
5.3	Primerjava dveh metod na testni množici SET3 z 227 obrazi.	91
6.1	Uporaba metode na osnovi barv brez in z uporabo heurističnih pravil na testni množici 472 barvnih slik, ki vsebujejo samo frontalne obraze.	101
6.2	Uspešnost kombinirane metode na testni množici slik.	106

Poglavje 1

Uvod

S povsod navzočo informacijsko in večpredstavno tehnologijo se razvijajo vse bolj napredne in uporabniku prijazne metode za komunikacijo z računalnikom, ki ne temeljijo na običajnih napravah, kot so tipkovnica, računalniška miška in zaslon, temveč na slikovnih informacijah, pridobljenih iz video naprav, priključenih na računalnike. K temu je pripomoglo tudi zmanjševanje razmerja med ceno in kakovostjo naprav za zajemanje videa. Sistemi za zajemanje videa so sedaj vse pogostejša oprema namiznih računalnikov.

Hiter razvoj na področju raziskovanja človeških obrazov nam obljublja, da bomo lahko iz slike dobili informacije o identiteti osebe, stanju osebe in s tem tudi vsaj delni vpogled v namen osebe. Računalniki se bodo lahko na podlagi teh informacij primerno odzvali. Ko bo računalnik na primer prepoznal izraz na obrazu osebe, bo ustrezno odreagiralo. Ravno na področju prepoznavanja človekovega izraza je bilo v zadnjih petih letih narejenih veliko raziskav in na podlagi tega je bilo izdelanih tudi veliko komercialnih izdelkov. V zadnjem času, ko so teroristične grožnje vse bolj prisotne, pa so sistemi za prepoznavo človeških obrazov še posebej pomembni. Vse več prostorov je neprestano nadzorovanih s pomočjo video kamer, ki iščejo na slikah obraze in jih primerjajo z obrazi iskanih oseb. Nekatere države so za vstop, kot dopolnilo pristopnim obrazcem, že uvedle sistem, ki identificira osebo na podlagi biometričnih podatkov, pridobljenih iz prstnega odtisa in slike obraza. Samo še vprašanje časa je, kdaj bomo morali tudi v bolj vsakodnevnih situacijah svojo identiteto dokazovati na podlagi biometričnih podatkov.

Prvi korak pri katerem koli sistemu za analiziranje obraza pa je poiskati obraz na sliki.

Samo iskanje obrazov na sliki je za raziskovalce pravi izziv zaradi raznolikosti v velikosti obrazov, različnih lokacij obrazov na slikah, različnih položajev (pokončen, vodoraven) ter različnih poz (od spredaj, od strani). Različni izrazi na obrazu, delno zakriti obrazi in različne osvetlitve pa iskanje še otežijo. Slika 1.1 prikazuje nekaj različnih primerov slik obrazov.



Slika 1.1: Različni primeri slik obrazov: a) obrazi v različnih pozah, b) različnih velikosti, c) z različnimi izrazi na obrazu, d) različno osvetljeni in e) obrazi različnih polti (vir: podatkovna baza FERET [49] in podatkovna baza Yale [20]).

Človeški obraz je eden najpomembnejših delov telesa, na podlagi katerega razločujemo ljudi med seboj. Tudi za človeka je lahko samo razločevanje zelo težka naloga. Ljudje med seboj zelo dobro ločimo obraze iste polti kot je naša, pri obrazih, ki imajo drugačno kožno polt, pa imamo občutek, da so si med seboj zelo podobni. Veliko problemov imamo tudi pri razlikovanju obrazov, če jih vidimo obrnjene narobe. Vsi ti primeri nakazujejo na to, kako težka je naloga prepoznavanja človeškega obraza tudi za človeka, še posebno kadar imamo opraviti z obrazi, s katerimi nimamo veliko izkušenj (npr. zaradi drugačne polti).

Z raziskovanjem prepoznavanja človeških obrazov na sliki je povezanih veliko nalog, in ker se bomo v tej magistrski nalogi ukvarjali predvsem z iskanjem obrazov na sliki, si podrobneje pogledimo definicijo te naloge.

Iskanje obrazov je postopek, s katerim določimo točen položaj in obseg obrazov na sliki.

Izzivi, ki so tesno povezani z iskanjem obrazov, so:

Poza: Obraz je lahko na sliki v različnih pozah (od spredaj, pod kotom 45 stopinj, od strani, itd.) (slika 1.1a). Težava, ki se pri tem pojavlja, je v tem, da so pri določenih pozah značilke obraza, kot so nos, obrvi in oči, delno ali v celoti prikrite.

Prisotnost obraznih dodatkov: Značilke obraza, kot so brada, brki in očala, so lahko prisotne ali pa ne. Še dodatno, te značilke so lahko zelo različnih oblik, barv in velikosti (slika 1.1b).

Mimika obraza: Videz obraza je zelo povezan s samo mimiko obraza. Obraz lahko izgleda zelo različno, če se oseba na sliki čudi, smeji ali je žalostna (slika 1.1c).

Prekrivanje: Obrazi so lahko delno zakriti z različnimi predmeti, kot so: modni dodatki, deli telesa, lasje. Na skupinski sliki pa se lahko obrazi med seboj prekrivajo.

Geometrija slike: Na sam izgled obrazov na sliki lahko vpliva tudi položaj kamere glede na optični center. Na pravilni izgled slike pa vplivajo tudi različne geometrijske napake leč.

Pogoji pri zajemanju slike: Ko se slika zajame, nanjo zelo vpliva osvetlitev (moč osvetlitve, spekter in porazdeljenost osvetlitve) (slika 1.1d). Na videz obrazov na zajeti sliki vplivajo tudi same karakteristike naprav, s katerimi zajemamo sliko (npr. občutljivost senzorja).

Različna kožna barva: Ljudje različnih ras imajo med seboj zelo različno barvo polti (slika 1.1e). Uporaba različnih ličil pa lahko kožno barvo še dodatno spremenijo.

S problemom iskanja obraza na sliki je povezanih veliko podobnih problemov, ki pa jih moramo med seboj razlikovati:

Lokalizacija obraza: Namen je določiti, kje točno se nahaja obraz na sliki. Gre za poenostavljen problem iskanja obrazov, kjer slika vsebuje natanko en obraz.

Iskanje značilk obraza: Ugotoviti je potrebno ali so značilke prisotne in njihovo lokacijo na obrazu. Značilke so lahko oči, nos, nosnici, obrvi, usta, lica, ušesa, itd. Tudi tukaj se predpostavlja, da je na sliki natanko en obraz.

Prepoznavanje obraza ali identifikacija obraza: Pri tem problemu gre za primerjavo vhodne slike obraza s slikami, ki so v podatkovni bazi. Kot rezultat dobimo vrednost ujemanja slike obraza s podobnimi obrazi iz baze.

Potrjevanje obraza ali verifikacija: Namen je potrditi, predhodno najdene kandidate za obraz, ali res ustrezajo določenemu obrazu ali ne.

Sledenje obrazu: Metoda, s katero v realnem času, neprekinjeno sledimo položaju obraza skozi zaporedje slik.

Prepoznavanje obrazne mimike: Prepoznavanje razpoloženja osebe (vesel, jezen, žalosten, itd.).

Iskanje obrazov na sliki je prvi korak pri izgradnji sistema, ki bo reševal zgoraj naštetih probleme. Omeniti je potrebno, da veliko raziskovalcev definira svoje metode kot metode za iskanje obrazov, čeprav njihove metode uspešno najdejo obraze samo na slikah z enim obrazom. Te metode, ki bi jih lahko uvrstili med metode za lokalizacijo obraza, je potrebno ločevati od metod za iskanje obrazov, ki iščejo na slikah več obrazov.

Izmed vseh metod za iskanje obrazov so največ pozornosti pritegnile metode, ki temeljijo na strojnem učenju, saj so se izkazale za zelo uspešne. Te metode so seveda zelo odvisne od učne množice, zato je zelo pomembna ustrezna izbira le-te. Vendar se pri takšnih metodah pojavi problem, kako posamezno metodo ustrezno ovrednotiti. Veliko člankov na temo iskanja obrazov primerja med seboj različne metode na podlagi števila uspešno in napačno najdenih obrazov. Pri takšnih metodah pa so poleg tega pomembni še drugi parametri, kot so: čas učenja, čas razvrščanja ter število slik, ki so potrebne za učenje. Primerjava med različnimi metodami je zelo težka, ker so posamezne metode preizkušene na različnih testnih množicah slik. Poleg tega pa nekateri raziskovalci različno definirajo pravilno in napačno najdene obraze. Potrebno je poenotiti izraze kot so:

- **Pravilno najdeni obrazi (ang. detection rate, True Positives, TP):** To je število obrazov, ki so bili med vsemi obrazi na sliki (ročno prešteti) uspešno najdeni. Obraz je označen kot pravilno najden samo, če se obseg najdenega obraza za določen odstotek prekriva z ročno označenim obrazom. Število najdenih obrazov

se največkrat predstavi procentualno, kot razmerje med številom najdenih obrazov in vsemi obrazy.

Pri iskanju obrazov se lahko pojavita dve vrsti napak:

- **Obrazi, ki niso najdeni (ang. False Negative, FN):** To merilo označuje tiste obraze, ki jih s pomočjo programa za iskanje obrazov nismo našli. Posledično imamo zaradi tega manjši delež pri najdenih obrazih.
- **Napačno najdeni obrazy (ang. false detections, False Positives, FP):** Program za iskanje obrazov najde obraz, čeprav tam obraza ni.

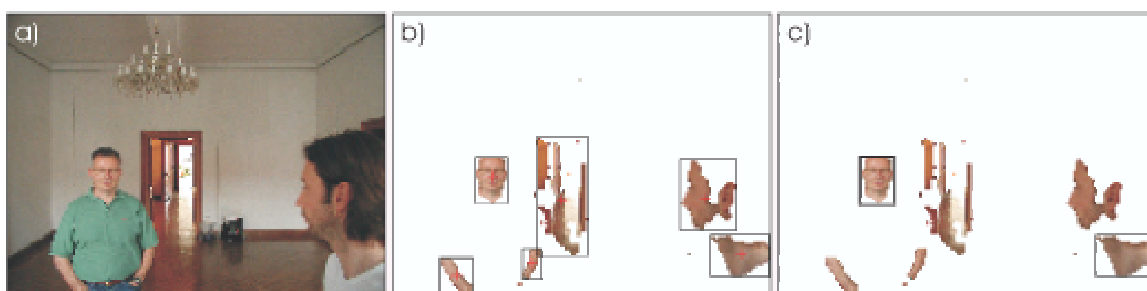
Po teh merilih bi se morali ravnati vsi raziskovalci, ki preizkušajo svoje metode za iskanje obrazov. Na ta način bi bile posamezne metode enostavnejše primerljive. Pomembno je tudi, da pri vsaki metodi navedemo vsa tri merila, saj si lahko posamezno metodo prilagodimo tako, da nam daje zelo dobre rezultate pri najdenih obrazih (TP) v zameno za visoko število napačno najdenih obrazov (FP).

	je obraz	ni obraza
našel obraz	TP (pravilno najdeni)	FP (napačno najdeni)
ni našel obraza	FN (obrazy niso najdeni)	TN (pravilno izločeni)

Tabela 1.1: Merila, ki se uporabljajo pri sistemih za iskanje obrazov na sliki.

Tabela 1.1 nam ponuja pregled opisanih meril, s katerimi se srečamo pri preizkušanju metod za iskanje obrazov. Skrajno levi stolpec označuje, ali je metoda našla obraz ali ne. Zgornja vrstica pa, ali je najdeni obraz dejansko obraz ali ne. Natančno moramo tudi določiti, kaj je za naš sistem obraz in kaj ni. Ali delno zakrit obraz še štejemo kot obraz? Ali so iskani obrazy omejeni glede na velikost, nagib ali obrat obraza. V tabeli 1.1 lahko opazimo tudi oznako TN (True Negative), ki označuje pravilno izločene neobraz. To merilo nam pove, koliko neobrazov je metoda v fazi potrditve pravilno izločila, da niso obrazy. Uporablja se kot vmesno merilo metode, za končni rezultat pa ni toliko pomembno. Vse obraze, ki jih iščemo na sliki, nam podaja seštevek: $TP+FN$.

Slika 1.2 nam prikazuje posamezna merila, na podlagi katerih ocenjujemo uspešnost iskanja obrazov na sliki. Prva slika nam prikazuje vhodno sliko, na kateri iščemo obraze. Na sliki imamo 2 obraze. Na drugi sliki je metoda odkrila 6 kandidatov za obraz na podlagi kožne barve. Na zadnji sliki je metoda na osnovi pravil ohranila samo 2 kandidata za obraz. Vidimo lahko, da je metoda uspešno izločila 3 kandidate za obraz, torej je $TN=3$. Med najdenimi obrazi je en pravilno najden, $TP=1$ (označen s povdarjeno črto) in en napačno najden, $FP=1$ (označen z nepovdarjeno črto). Metoda je napačno izločila oziroma ni našla enega obraza, $FN=1$. Iz slike lahko preverimo, da je seštevek $TP+FN$ enak številu vseh obrazov, ker velja $TP+FN=2$.



Slika 1.2: Merila pri iskanju obrazov: a) vhodna slika, b) izbrani kandidati za obraz, c) najdeni obrazi: pravilno najden obraz, napačno najden obraz in obraz, ki ni bil najden.

Raziskave na področju človeških obrazov so v računalniškemu vidu tudi zelo razširile področje prepoznavanja objektov. Pri prepoznavanju objektov je večina raziskav, ki temeljijo na modelih ali vizualnem izgledu, pa tudi na 3D modeliranju, omejena na toge predmete, ki jih opazujemo pod različnimi koti in pod različnimi osvetlitvami. Na problem iskanja obrazov lahko gledamo kot na problem prepoznavanja objektov z dvema možnima razredoma: izbrani del slike je lahko obraz in tako pripada razredu "obrazov", ali pa ni obraz in pripada razredu "neobrazov". Klasifikacijske metode, ki iz slike razberejo določeno informacijo, lahko velikokrat uporabljamo, poleg iskanja obrazov, tudi za iskanje drugih objektov na sliki. Lahko jih uporabimo tudi pri klasifikaciji vzorcev ali pri postopkih učenja.

Če za vhod učnega algoritma uporabimo vse slikovne elemente slik, bo dimenzija takšne učne množice izjemno velika. Takšno učno množico bodo namreč sestavljali vsi

slikovni elementi slike. Zaradi tega se uporabljajo različne tehnike, ki iz učnih primerov slik obrazov in neobrazov izločijo samo karakteristične slikovne elemente, ki sliko zelo dobro opisujejo. Tako opisane slike so največkrat predstavljene z večdimenzionalno porazdelitveno funkcijo, katere robni pogoji so večinoma nelinearne funkcije. Da dosežemo ustrezno učinkovitost pa morajo že sami klasifikatorji iz učne množice izbirati ustrezne učne primere.

V nadaljevanju bom naredil pregled različnih tehnik za iskanje obrazov na sliki. Spoznali bomo različne metode, ki jih bom smiselno združil v štiri skupine. Ker se bom v magistrskem delu ukvarjal z iskanjem obrazov na barvnih slikah, je zelo pomembno, da najprej spoznamo različne barvne prostore in metode za segmentacijo barv, čemur bo namenjeno tretje poglavje. V četrtem poglavju bom s pomočjo različnih metod poskušali segmentirati kožno barvo v izbranem barvnem prostoru. V petem poglavju bomo spoznali metodo BDF, ki na osnovi naučenih modelov obraza in neobraza s pomočjo Bayesovega klasifikatorja išče frontalne obraze na sivinski sliki. Glavni prispevek magistrske naloge pa je predstavljen v šestem poglavju, kjer predstavim izboljšano metodo za iskanje frontalnih obrazov na barvnih slikah, ki kombinira metodo na osnovi barv in metodo BDF. Predlagano metodo tudi preizkusim na testni množici barvnih slik, ki vsebujejo frontalne obraze. Na koncu združim ugotovitve, predlagam izboljšave nove metode ter izpostavim poudarke in smernice za nadaljno raziskovanje.

Poglavje 2

Pregled različnih tehnik za iskanje obrazov na sliki

Za iskanje človeških obrazov na slikah obstaja ogromno različnih tehnik, med katerimi je zelo težko izbrati najpomembnejše in jih ustrezno združiti v tematske skupine. Tudi člankov na to temo je ogromno, kar še poveča težavnost izbire. V tem poglavju bomo poslušali narediti pregled nad obstoječimi tehnikami za iskanje obrazov na sivinskih in barvnih slikah. Čeprav nekatere metode uporabljajo različne pristope, jih vseeno lahko razdelimo v štiri glavne skupine [81]:

1. **Metode na osnovi znanja (ang. Knowledge-based methods):** Te metode vsebujejo pravila, ki opisujejo človekovo poznavanje tipičnih karakteristik obraza. Običajno pravila opisujejo povezave med posameznimi značilkami obraza. Te metode se običajno uporablja za lokalizacijo obraza.
2. **Metode na osnovi nespremenljivih značilk (ang. Feature invariant approaches):** Cilj teh algoritmov je poiskati strukturne značilke obraza, ki se ne spremenijo, če jih opazujemo pod različnimi koti, pozami ali pod različnimi osvetlitvami. Na podlagi teh značilk nato iščemo obraz. Tudi te metode se uporabljajo predvsem pri določitvi lokacije obraza.
3. **Metode ujemanja s šablono (ang. Template matching methods):** Obstaja kar nekaj standardnih vzorcev ali šablon, ki opisujejo obraz ali posamezen del obraza

oziroma značilko obraza. Za iskanje obraza na sliki se izračuna ujemanje med vhodno sliko in shranjenim vzorcem. Te metode se uporabljajo tako za lokalizacijo kot tudi za iskanje obrazov.

4. **Metode na osnovi videza (ang. Appearance-based methods):** V primerjavi s prejšnjo metodo se tukaj model ali vzorec določi s pomočjo učne množice, ki reprezentativno določa posamezne obrazne značilnosti. Tako naučeni modeli nato služijo za iskanje obrazov.

V posamezno, zgoraj opisano, skupino uvrščamo veliko metod. V nadaljevanju bomo podrobneje opisali glavne predstavnike posamezne skupine, sedaj pa naštejmo vse metode, ki se uporabljajo za iskanje obrazov na sliki, po posameznih skupinah:

Pristop na osnovi znanja (od zgoraj navzdol):

- metode na osnovi pravil nad različno velikimi slikami.

Pristop na osnovi značilk (od spodaj navzgor):

- uporaba značilk obraza,
- uporaba teksture obraza,
- uporaba barve kože,
- kombinacija več značilk.

Pristop na osnovi šablone:

- uporaba pred-pripravljenih šablon,
- uporaba spremenljivih šablon.

Pristop na osnovi videza:

- razčlenitev na podlagi lastnih vektorjev oziroma uporaba lastnih-obrazov (ang. Eigenfaces),
- uporaba porazdelitvenih modelov,

- uporaba nevronske mreže,
- metoda podpornih vektorjev (SVM),
- Naivni Bayesov klasifikator,
- skriti Markov model (HMM),
- pristop na osnovi teoretičnih informacij (ang. Information-theoretical approach),
- uporaba induktivnega učenja.

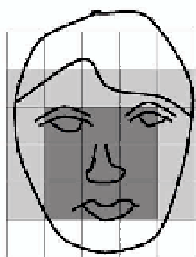
2.1 Pristop na osnovi znanja - od zgoraj navzdol

Pri temu načinu metode za iskanje obrazov temeljijo na človeškem znanju o značilnostih obraza. Na osnovnih značilnostih obraza lahko enostavno določimo nekatera pravila, na podlagi katerih nato določimo značilke obraza in njihovo medsebojno odvisnost. Na primer: obraz na sliki velikokrat vsebuje dvojce simetričnih oči, nos in usta. Medsebojno odvisnost značilk lahko predstavimo s pomočjo njihovega položaja ali pa medsebojne razdalje. Iz vhodne slike najprej izločimo značilke, nato pa kandidate za obraz preverimo s pomočjo določenih pravil. To preverjanje je namenjeno predvsem izločanju napačno najdenih obrazov.

Problem metod na osnovi znanja se pojavi pri prenosu znanj o lastnostih obraza v dobro definirana pravila. Če so pravila preveč natančna, lahko kakšen obraz izpustimo, ker ne izpolnjuje vseh pravil. Če pa so pravila preveč splošna, lahko dobimo preveč napačnih zadetkov (ang. False Positives, FP). Težava je tudi v tem, da takšne metode zelo težko razširimo, da bi našle obraze tudi v različnih položajih, dokler seveda ne podamo pravil za vse položaje. Takšne hevristične metode najbolje delujejo na frontalnih obrazih, ki so na enoličnem ozadju.

Yang in Huang [76] sta uporabila hierarhično metodo, ki temelji na osnovi znanja. Njun sistem uporablja tri nivoje pravil. Na najvišjem nivoju pregledata celotno sliko in na ta način najdeta kandidate za obraz. Na vsakem nivoju nižje definirata podrobnejšo obliko obraza. Na najnižjem nivoju so tako opisane posamezne značilke obraza. Hierarhija se gradi s pomočjo zmanjševanja velikosti slike. Pri najnižji velikosti oziroma kvaliteti slike se išče obraz na podlagi sledečih pravil (Slika 2.1):

1. središče obraza vsebuje 4 slikovne elemente z enotno intenzivnostjo,
2. ti slikovni elementi so obkroženi s slikovnimi elementi enake intenzivnosti in
3. razlika v intenzivnosti med sredinskimi in slikovnimi elementi na robu je velika.



Slika 2.1: Primer obraza, ki se uporablja pri pristopu na osnovi znanja: pravila so definirana na osnovi intenzivnosti slikovnih elementov obraznih značilk.

Kandidati, izbrani na tem nivoju, se nato preverijo na naslednjem nivoju, kjer uporabimo pravila na sliki boljše kvalitete. Na tem nivoju se na izbranih kandidatih za obraz uporabi lokalni histogram in metoda za iskanje robov. Kandidati, ki ustrezajo pravilom na drugem nivoju se nato preverijo še na tretjem nivoju, na katerem se iščejo značilke, kot so oči in usta. Rezultati so pokazali, da je na bazi 60 slik, sistem uspešno lociral 50 obrazov z 28 napačno najdenimi obrazi. Čeprav metoda ne daje preveč dobrih rezultatov, je tukaj zelo pomemben pristop, ki uporablja hierarhijo različnih velikosti slike in stopnjevanje pravil.

Kotropoulos in Pitas [35] sta predstavila metodo, ki se od prej opisane metode razlikuje v tem, da na prvem nivoju uporablja projekcijsko metodo. Naj bo $I(x, y)$ vrednost slikovnega elementa na položaju (x, y) slike velikosti $m \times n$. Horizontalna in vertikalna projekcija slike je potem definirana kot: $HI(x) = \sum_{y=1}^n I(x, y)$ in $VI(y) = \sum_{x=1}^m I(x, y)$. Najprej se izračuna horizontalni profil slike, na katerem določata lokalna minimuma levo in desno stran glave. Vertikalni profil pa na podlagi lokalnih minimumov določa položaj obraznih značilk, kot so: oči, nos in usta. To metodo je uporabil tudi Kanade [31] za uspešno odkrivanje robov obraza.

Metoda je bila preizkušena na ACTS M2VTS (MultiModal Verification for Teleservices and Security applications) podatkovni bazi [51], ki vsebuje video izseke 37 različnih ljudi. Vsak posnetek vsebuje samo en obraz na enotnem ozadju. Metoda uspešno odkrije 86.5% vseh obrazov. Metoda potrdi samo tiste kandidate za obraz, pri katerih so najdene vse značilke obraza. Metoda je neprimerna za šumno ozadje in v primeru, ko slika vsebuje več obrazov. Za uspešnost metode je pomembna tudi pravilna izbira področja, nad katerim se vrši iskanje.

2.2 Pristop na osnovi značilke - od spodaj navzgor

V primerjavi z metodami, ki temeljijo na znanju ali s tako imenovanimi metodami od zgoraj navzdol, so nekateri raziskovalci poizkusili s preiskovanjem slike pri dnu, oziroma s tako imenovanimi metodami od spodaj navzgor. Pri teh metodah se na sliki najprej poišče takšne značilke obraza, ki se ne spreminjajo. Izhajali so namreč iz ugotovitve, da človek lahko enostavno prepozna na sliki obraz in objekte v različnih položajih in osvetlitvah, torej morajo obstajati takšne lastnosti ali značilke, ki se pod temi pogoji ne spreminjajo. Veliko metod predlaga, da se najprej poiščejo značilke na podlagi katerih nato sklepamo, ali je to obraz ali ne. Značilke obraza, kot so obrvi, oči, nos, usta in lasje, se večinoma išče s pomočjo metod za iskanje robov. Na podlagi tako najdenih značilk se nato zgradi statistični model, ki opiše njihovo medsebojno odvisnost in potrdi, ali značilke določajo obraz. Problem pri teh metodah je, da so značilke velikokrat zakrite z drugimi objekti ali slabo vidne zaradi različnih osvetlitev ali šuma na sliki. V nekaterih primerih so meje med značilkami zelo slabe, po drugi strani pa sence povzročajo dodatne robove. Vse te težave skupaj imajo za posledico manjšo uporabnost takšnih algoritmov. V nadaljevanju si bomo pogledali posamezne metode, ki uporabljajo različne značilke oziroma njihovo kombinacijo za odkrivanje obrazov na slikah.

2.2.1 Metode na osnovi značilke obrazov

Sirohey [61] je predlagal lokalizacijsko metodo za segmentacijo obraza iz šumnega ozadja. Za odstranitev in združitev robov je uporabil metodo za iskanje robov in hevristiko, tako da so se ohranili samo robovi na obrazu. Obrisu obraza je nato prilagodil elipso, s katero

je ločil obraz od ozadja. Uspešnost tega algoritma je 80% na podatkovni bazi z 48 slikami na šumnem ozadju.

Chetverikov in Lerch [12] sta predstavila enostavno metodo iskanja obrazov s pomočjo regij (ang. blobs) in črt, sestavljenih iz zaporedja isto usmerjenih robov. Njuni modeli so vsebovali dve temni in tri svetle regije, ki so predstavljale oči, ličnici in nos. Črte pa so predstavljale obris obraza, obrvi in lica. Na podlagi dveh trikotnih kombinacij sta nato iskala ustrezne regije, ki pripadajo obrazu. Tako izbrane regije, ki so jih omejevale črte, so bile nato potrjene kot kandidati za obraz.

Graf in sodelavci [23] so razvili metodo, ki določi položaj značilk in obraza na sivinskih slikah. Po filtriranju s pasovnim (ang. "band-pass") filtrom so uporabili morfološke (ang. morphological) operacije, da so našli regije z visoko intenziteto določenih oblik, kot so na primer: oči. Histogram slike je pokazal izrazito konico, na podlagi katere so določili mejno vrednost za binarizacijo slike. Na binarni sliki so določili regije značilk in njihovo kombinacijo potrdili s klasifikatorjem. Metoda je bila preizkušena nad slikami 40 obrazov in s pomočjo 5 video posnetkov, ki so vsebovali od 100 do 200 slik. Ni pa jasno, kako so bile uporabljene morfološke operacije in kako so bile značilke kandidatov upoštevane pri iskanju obraza.

Leung in sodelavci [38] so izdelali metodo, ki temelji na verjetnosti, za določitev pozicije obraza na šumnem ozadju. Metoda je osnovana na metodah za iskanje lokalnih značilk in na preiskovanju naključnih grafov (ang. random graph matching). Njihov namen je bil predstaviti problem lokalizacije obraza kot problem ujemanja naključnih grafov. Cilj preiskovanja grafov je najti ustrezno postavitve značilk obraza, ki predstavljajo vzorec obraza. Za opis obraza so uporabili pet obraznih značilk, kot so: obe očesi, nosnici in spoj med nosom in ustnicami. Med istimi pari značilk so izračunali relativno razdaljo in določili Gaussovo porazdelitev razdalj na vseh slikah. Končni obrazni vzorec so določili s pomočjo rezultatov različno obrnjenih in različno velikih slik, predhodno obdelanih z Gaussovim filtrom. Značilke obraza so nato določili s pomočjo primerjave slikovnih elementov z obraznim vzorcem. Izbrali so dve najboljši značilki in na osnovi teh so nato poiskali ostale. S pomočjo statističnega modela razdalj so nato določili najboljše razporeditve značilk. Končni kandidati značilk so bili tako sestavljeni samo iz značilk, ki so najboljše ustrezale razdaljam, določenimi z modelom. Na ta način so določili najboljšo kombinacijo značilk,

ki je najverjetneje določala obraz. Iskanje najboljše kombinacije značilk so predstavili kot problem ujemanja naključnih grafov. Vozlišča grafa so predstavljale značilke, povezave med njimi pa razdalje med značilkami. Verjetnost pravilne postavitve značilk so nato preverili s pomočjo verjetnostne funkcije. Za preizkus so uporabili množico 150 slik, na katerih so predpostavili, da je obraz pravilno določen, če je sistem našel tri značilke v pravem razmerju. Takšen sistem je uspešno našel 86% vseh obrazov.

Leung in sodelavci [9, 39] so kasneje namesto medsebojnih razdalj med značilkami uporabili alternativno metodo, ki je temeljila na statistični teoriji oblike (ang. statistical theory of shape), ki so jo razvili Kendall [32] ter Mardia in Dryden [41]. Ta metoda združuje verjetnostno funkcijo gostote nad N točkami značilk. Prednost te metode je v tem, da odkrije tudi delno zakrite obraze. Ni pa znano, ali uspešno najde tudi več obrazov na sliki.

Yow in Cipolla [82, 83] sta predstavila metodo na osnovi značilk, ki uporablja veliko količino podatkov iz same slike. Na vhodni sliki najprej uporabita drugi odvod Gaussovega filtra. Zanimive točke, ki se pokažejo na lokalnih maksimumih po implementaciji filtra, predstavljajo možne kandidate za značilke. Okrog zanimivih točk nato poiščeta robove in jih združita v regije. Robove združujeta na osnovi njihove podobnosti v orientaciji in dolžini. Izračunane lastnosti posameznih regij, kot so dolžina robov, poudarjenost robov in varianca intenzivnosti slikovnih elementov, tvorijo vektor značilk. Iz učne množice obraznih značilk nato izračunata še povprečno vrednost in kovariančno matriko posameznega vektorja značilk. Izbrani del slike je označen za obrazno značilko, če je Mahalanobisova razdalja od tako določenega vektorja značilk pod mejno vrednostjo. Označene značilke nato združita na podlagi modela znanj, da so med seboj v pravilnem razmerju. Tako združene značilke nato preverita še s pomočjo Bayesove mreže. Opazna prednost te metode je v tem, da najde obraze različnih orientacij in poz. Uspešnost metode, preizkušene na 110 slikah obrazov različnih velikosti, orientacij in iz različnih pogledov, je bila 85%. Neuspešna je bila v 28% in delovala je samo na obrazih večjih od 60×60 slikovnih elementov. Naknadno pa je bila metoda dopolnjena še z modeli aktivnih obrisov (ang. Active contour models).

Han in sodelavci [24] so razvili tehniko na osnovi morfologije za učinkovito iskanje tako imenovanih "očem analognih" segmentov (ang. eye-analogue segments). Trdijo, da so

oči in obrvi najbolj poudarjene in nespremenljive značilke obraza. Zaradi tega, določijo za "očem analogne" segmente prav robove, ki so oblikovani v obliki oči. Segmente, ki predstavljajo oči izločijo s pomočjo morfoloških operacij, na podlagi katerih nato iščejo ostale značilke obraza, ki so v pravem razmerju z očmi. Najdene kandidate na koncu še preverijo s pomočjo nevronske mreže. Njihovi preizkusi kažejo 94% točnost na testni množici 122 slik s 130 obrazi.

Amit in sodelavci [3] so predstavili metodo za iskanje oblike, ki so jo uporabili za iskanje frontalnih obrazov na slikah. Najprej so se osredotočili na porazdeljenost robov, ki je značilna samo za obraze. S pomočjo induktivne metode so določili veliko različnih kombinacij robov, ki so neodvisne od različnih transformacij slike in so bolj značilne za obraze kot za ozadje. S pomočjo algoritma CART [6] so zgradili klasifikacijsko drevo, na podlagi katerega so nato določili ali je opazovani del slike obraz ali ne. Njihovi rezultati na testni množici 100 slik iz Olivetti podatkovne baze [58] prikazujejo napake, in sicer za FP 0.2% in za FN 10%.

2.2.2 Metode na osnovi teksture obraza

Človeški obraz ima določeno teksturo, na podlagi katere ga lahko razlikujemo od drugih objektov. Augusteijn in Skufca [4] sta razvila metodo, ki sklepa na prisotnost obraza na sliki na osnovi obrazu podobne teksture. Teksturo izračunata s pomočjo statističnih značilk drugega reda (ang. second-order statistical features, SGLD) v slikah velikosti 16×16 slikovnih elementov. Upoštevani so trije tipi značilk, kot so koža, lasje in ostale značilke. Za klasifikacijo sta uporabila nevronske mreže, za tvorbo različnih razredov posameznih značilk pa Kohonenovo samoorganizacijsko mrežo značilk (ang. Kohonen self-organizing feature map). Glede na točnost teksture kože in las sta določila, ali pripada določena tekstura obrazu. Kot rezultat sta podala samo vrednost pravilne klasifikacije tekstur. Metode pa nista preizkusila za lokalizacijo ali iskanje obrazov.

Dai in Nakano [18] sta predlagano metodo SGLD uporabila tudi za iskanje obrazov. Kot značilko obraza sta uporabila kožno barvo. S pomočjo obrazne teksture sta izdelala model iskanja obrazov na barvnih slikah, s pomočjo katerega sta označila vse barvne regije, ki so vsebovale obraze. Prednost takšnega pristopa je v tem, da s to metodo najdemo tudi obraze, ki niso nujno pokončni in vsebujejo tudi druge obrazne dodatke, kot so brada

ali očala. Metoda se je zelo dobro izkazala na preizkusu nad testno množico 30 slik, ki so vsebovale 60 obrazov.

2.2.3 Metode na osnovi barve kože

Kožna barva se je izkazala za zelo učinkovito značilko, ki se lahko uporablja za iskanje obrazov ali za sledenje drugim delom telesa, kot so na primer roke. Čeprav imajo različni ljudje različno barvo kože, je veliko študij dokazalo, da je glavna razlika med njimi v njihovi svetlosti in ne toliko v njihovi barvi oziroma kromatičnosti barve (ang. chrominance). Za opis kožne barve je bilo uporabljenih že veliko barvnih prostorov: barvni prostor RGB, normalizirani prostor RGB, barvni prostor HSV oziroma HSI, YCrCb, YIQ, YES, CIE XYZ in CIE LUV. Posamezni barvni prostori so podrobneje opisani v 3. poglavju. Predlaganih je bilo tudi veliko različnih tehnik za izgradnjo barvnega modela. Najenostavneje lahko model določimo s pomočjo mejnih vrednosti ali na podlagi pravil. Za kromatični barvni prostor CrCb, določimo na primer mejni vrednosti $(Cr1, Cr2)$ in $(Cb1, Cb2)$. Opažovani slikovni element z barvnima komponentoma Cr in Cb torej pripada kožni barvi, če velja: $Cr1 \leq Cr \leq Cr2$ in $Cb1 \leq Cb \leq Cb2$.

Crowley in Coutaz [15, 16] sta kot barvni model uporabila histogram $h(r, g)$ v normaliziranem prostoru RGB. Posamezna barva je predstavljena s pomočjo dveh vrednosti r in g . Barvni element klasificiramo kot kožni, če je $h(r, g) \geq \tau$. Vrednost τ označuje mejno vrednost, ki je izbrana na podlagi histograma učne množice. Kot nasprotje ne-parametričnim modelom, za katere se uporablja histogram, se uporabljajo tudi Gaussovi modeli oziroma mešanice Gaussovih modelov (ang. mixture of Gaussians). Parametri enojnih Gaussovih modelov se določijo s pomočjo metode največje verjetnosti (ang. maximum-likelihood). Mešanico Gaussovih modelov se uporablja v primeru, če kožna barva različnih etičnih skupin tvori več različnih skupkov in ne samo enega samega, tako kot pri enojnem Gaussovem modelu. Parametre za takšen model pa običajno določimo s pomočjo EM algoritma [26, 79].

Jones in Rehg [29] sta izvedla preizkus dveh različnih tehnik nad približno enim bilijonom kožni barvi podobnih slikovnih elementov v barvnem prostoru RGB. Primerjala sta mešanico Gaussovih modelov in model na osnovi histograma. Na podlagi rezultatov sta zaključila, da je metoda na osnovi histograma učinkovitejša.

Barvna informacija je zelo učinkovito orodje za iskanje barvnih regij, kjer se nahaja obraz ali za določevanja barvnih regij obraznih značilk. Ta informacija pa je uporabna samo, če uspešno najdemo barvo tudi pod različnimi svetlobnimi pogoji. Izkaže se namreč, da je označevanje barv neučinkovito, če se osvetlitev na sliki zelo močno spreminja, kajti videz barve se pod različno osvetlitvijo zelo spreminja. Zaradi tega pridejo v takšnih primerih v poštev metode, ki ohranjajo barvno konstantnost.

McKenna in sodelavci [44] so predstavili prilagodljiv barvni model za sledenje obrazu pod spremenljivimi svetlobnimi pogoji. Namesto, da se bi zanašali samo na barvni model kože, ki temelji na konstantnosti barve, so uporabili stohastičen model, ki se na podlagi nenehnega ocenjevanja barvne porazdelitve prilagaja spremembam. Rezultati kažejo, da lahko sistem uspešno sledi obrazu pod različnimi pogoji osvetlitve. Seveda pa predlagane metode ne moremo uporabiti na mirujočih slikah.

Za iskanje obrazov na sliki pa samo barva kože ne zadošča, zato se metode na osnovi barv velikokrat uporablja v kombinaciji z drugimi metodami, ki temeljijo na drugih značilkah. Nekaj takšnih sistemov bomo opisali v nadaljevanju.

2.2.4 Metode na osnovi kombinacije različnih značilk

Predlaganih je bilo veliko metod, ki kombinirajo med seboj več obraznih značilk. Večina s pomočjo globalnih značilk (kot so barva kože, velikost in oblika obraza), najprej poišče kandidate za obraz. Za potrditev kandidatov pa uporabijo lokalne, bolj podrobne značilke kot so oči, obrvi, nos ali lasje. Tipičen pristop uporabi najprej koži podobno barvo, da ohrani samo kožno barvo. Nato s pomočjo metode "rast regij" določi regije. Če je oblika tako združenih regij ovalna, postanejo regije kandidati za obraz. Za potrditev kandidatov za obraz se nato uporabi še lokalne značilke. Poenostavljen pristop uporablja tudi metoda za iskanje obrazov pri interaktivni instalaciji "15 sekund slave" [63].

Yachida in sodelavci [11, 75, 74] so predstavili metodo za iskanje obrazov na barvnih slikah, ki temelji na mehki (ang. fuzzy) logiki. Uporabili so dva mehka modela za predstavitev porazdeljenosti barve kože in barve las v CIE XYZ barvnem prostoru. Za predstavitev obraza na sliki so uporabili pet različnih modelov: en model za frontalni položaj obraza in 4 modele za pogled od strani. Vsak model je sestavljen iz $m \times n$ celic, ki vsebujejo več slikovnih elementov. Vsaka celica je opisana s pomočjo dveh atributov, ki določata

razmerji med kožno barvo in barvo las, glede na preostalo barvo celic. Na testni sliki je lahko vsak slikovni element klasificiran kot obraz, lasje, lasje/obraz ali lasje/ozadje, na podlagi porazdelitvenih modelov, ki določata lase in obraz. Tako označene regije obrazov in las se nato primerjajo še z obrazu podobnimi modeli in če nato ustrezajo modelu, so izbrane za kandidate za obraz. S pomočjo horizontalnih robov so preverili še pare značilk, kot so oči in obrvi ter nos in usta, in s tem potrdili, da so kandidati res obrazi.

Sobottka in Pitas [62] sta predlagala metodo za lokalizacijo obraza in izločitev obraznih značilk na osnovi oblike in barve. Najprej sta sliko barvno segmentirala v HSV barvnem prostoru. Na ta način sta na sliki ohranila samo koži podobne regije, ki sta jih s pomočjo metode "rast regij" združila. Združevanje regij sta izvedla na pomanjšani sliki. Tako izločeno regijo sta predlagala za kandidata za obraz, če se je ustrezno prilagodila elipsi. Kandidate sta nato potrdila na podlagi najdenih značilk znotraj elipse. Pri iskanju značilk, kot sta oči in usta, sta izhajala iz predpostavke, da sta temnejše barve kot obraz.

Terrillon in sodelavci [67, 68] so za barvno segmentacijo uporabili Gaussov model. Za določitev pravilne oblike obraza pa so izračunali 11 osnovnih geometrijskih momentov, s pomočjo katerih so nato naučili nevronske mreže. Preizkuse so izvedli na bazi 100 slik in dosegli 85% točnost.

2.3 Pristop na osnovi šablone

Šablona v obliki obraza, ki se uporablja za iskanje obraza, je večinoma pred-pripravljena ali vnaprej določena s pomočjo parametrov, ki določajo primerjalne funkcije. Sliko primerjamo s šablono tako, da izračunamo stopnjo ujemanja vzorca z vsakim delčkom slike posebej. Tudi za posamezni obrazni vzorec, kot je vzorec oči, nosa in ust, izračunamo vrednosti ujemanja, na podlagi katerih nato določimo obraz. Prednost tega pristopa je v enostavni realizaciji. Dokazano pa je, da s pomočjo takšnega pristopa ne moremo najti obrazov različnih poz, velikosti in oblik. Predlagane so bile različne tehnike, ki s spremenljivimi šablonami na različnih velikostih slik poskušajo doseči neodvisnost metode od velikosti in oblike iskanega obraza.

2.3.1 Pred-pripravljene šablone

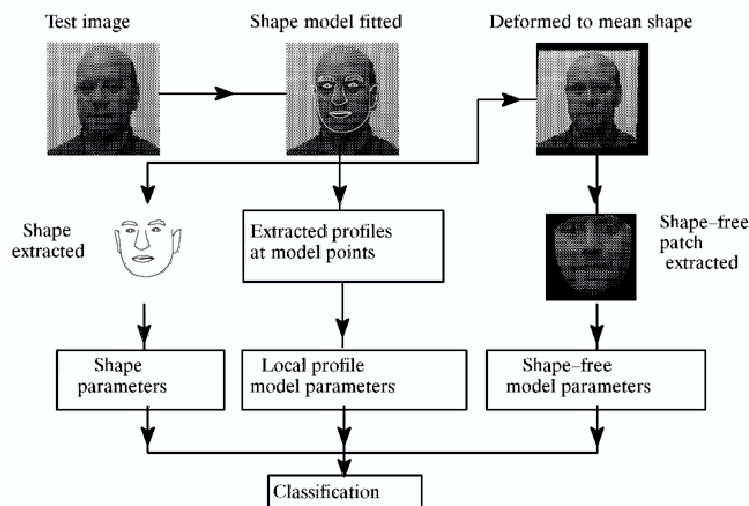
O enem od prvih poskusov iskanja frontalnih obrazov na fotografijah je poročal Sakai s sodelavci [57]. Za modeliranje obraza so uporabili različne podšablone za oči, nos, usta in obris obraza. Vsaka od teh šablon je bila predstavljena kot odsek linije. Linije so iz vhodne slike izločili na podlagi največje spremembe njihovega naklona in jih nato primerjali s podšablonami. V prvem koraku avtorji izračunajo korelacije med podslikami in šablono obrisa obraza, s čimer določijo možne lokacije obrazov. Nato na izbranih področjih slike izvedejo še primerjavo z ostalimi podšablonami. Povedano z drugimi besedami: v prvem koraku določimo središče pozornosti oziroma področje zanimanja, v drugem koraku pa poiščemo podrobnosti, da ugotovimo prisotnost obraza. Ta ideja, da bi obraze iskali s pomočjo podšablon samo na izbranih zanimivih področjih, je bila kasneje še velikokrat uporabljena.

Craw in sodelavci [13] so predstavili metodo lokalizacije, ki temelji na šabloni frontalne oblike obraza. Za izločanje robov so uporabili Sobelov filter. Robove so nato združili in jih primerjali z obliko obraza. Ko so na sliki našli obliko obraza, so postopek ponovili pri različnih velikostih slike, z namenom poiskati posamezne značilke, kot so oči, obrvi in usta. Kasneje so Craw in sodelavci [14] opisali metodo lokalizacije obrazov z uporabo 40 šablon za iskanje značilk obrazov, ter kontrolne strategije ocenjevanja rezultatov, najdenih s pomočjo na šablonah temelječih iskalnikov značilk.

2.3.2 Spremenljive šablone

Lenitis in sodelavci [36] so opisali obraz s pomočjo modela, ki je upošteval obliko obraza in intenzivnost slikovnih točk na obrazu. Na posameznih slikah iz učne množice so ročno označili obrise značilk, kot so obris oči, nosa, ličnic. Nato so posamezne točke obrisov predstavili s pomočjo vektorja. Za predstavitev vseh vektorjev oblik so uporabili model porazdelitve točk (ang. point distribution model, PDM). Za opis videza obraza na osnovi intenzivnosti slikovnih točk so uporabili podoben pristop, kot sta ga uporabila že Kirby in Sirovich [33]. S pomočjo modelov točk in modela aktivnih oblik (ang. active shape model, ASM) so določili lokacijo obraza in izračunali parametre oblike. Obraz so nato pretvorili v običajno obliko in izločili parametre intenzivnosti. Za klasifikacijo so uporabili obliko

obraza in parametre intenzivnosti. Na sliki 2.2 je prikazan postopek klasifikacije testne slike obraza.



Slika 2.2: Postopek klasifikacije, ki jo je uporabil Lenitis s sodelavci [36]. Testna slika se klasificira na podlagi modela oblike in izračunanih parametrov oblike ter parametrov intenzivnosti slikovnih točk.

2.4 Pristop na osnovi videza

V nasprotju s pristopom, ki temelji na šabloni, kjer obrazno šablono predhodno določi ekspert, tukaj vzorec ali šablono določimo sami s pomočjo učnih slik.

V glavnem imamo pri pristopu na osnovi videza opraviti s statistično analizo in strojnim učenjem, na podlagi katerega določimo glavne karakteristike slik obrazov in slik neobrazov. Naučene karakteristike predstavimo s pomočjo modelov porazdelitve ali diskriminantnih funkcij (ang. discriminant functions). Število karakteristik pogosto zmanjšamo z namenom večje učinkovitosti učenja in izboljšanja delovanja metode.

Veliko metod, ki delujejo na osnovi videza, bi lahko uvrstili med verjetnostne metode. Na vektor značilk, ki ga dobimo iz slike, namreč lahko gledamo kot na naključno spremenljivko \mathbf{x} . To naključno spremenljivko \mathbf{x} lahko označimo kot obraz ali kot neobraz na

podlagi verjetnostne porazdelitvene funkcije $p(\mathbf{x}|face)$ in $p(\mathbf{x}|nonface)$. Za klasifikacijo, izbrane slike na podlagi izračunane verjetnosti lahko uporabimo Bayesov klasifikator ali metodo maksimalne verjetnosti (ang. maximum likelihood). Bayesovega klasifikatorja ne moremo neposredno uporabiti iz več razlogov: dimenzija spremenljivke \mathbf{x} je zelo velika, verjetnostni funkciji $p(\mathbf{x}|face)$ in $p(\mathbf{x}|nonface)$ pa predstavljata kompleksen model.

Ostale metode, ki uporabljajo za iskanje obrazov pristop na osnovi videza, pa poskušajo najti ustrezno diskriminantno funkcijo (t.i. odločitveno ravnino, razdelitveno hiperravnino ali mejno funkcijo), ki bo razlikovala med obrazi in neobrazi. Običajno se diskriminantno funkcijo poišče tako, da se vzorce slik preslika v prostor z manjšo dimenzijo in na podlagi tega prostora se izoblikuje mejno funkcijo. Drugi način pa je, da s pomočjo nevronske mreže poiščemo ustrezno nelinearno odločitveno ravnino. Pri iskanju ustrezne ločnice si lahko pomagamo tudi s pomočjo metode podpornih vektorjev. Ta metoda deluje tako, da vzorce slik predstavi v višjedimenzijskem prostoru in na podlagi tega določi ravnino, ki ločuje med obrazi in neobrazi.

2.4.1 Lastni obrazi (ang. Eigenfaces)

Prvi, ki je uporabil metodo lastnih vektorjev na slikah obrazov je bil Kohonen [34]. Metodo za prepoznavo normaliziranih obrazov je predstavil s pomočjo nevronske mreže. Obraz je opisal s pomočjo lastnih vektorjev korelacijske matrike posameznih slik obrazov. Tako izračunane lastne vektorje slik obrazov so kasneje poimenovali lastni obrazi (ang. eigenfaces).

Kirby in Sirovich [33] sta pokazala, da lahko množico različnih slik obrazov predstavimo z zelo majhnim številom linearno odvisnih slik. Za takšno predstavitev slik sta uporabila analizo glavnih komponent (ang. principal component analysis) [27], imenovano tudi transformacija Karhunen-Loeve ali transformacija Hotelling. Idejo je prvič predstavil že Pearson leta 1901 in kasneje, leta 1993 še Hotelling.

Vsako sliko dimenzije $m \times n$ predstavimo kot vektor dolžine $m \cdot n$. Bistvo ideje je v tem, da poiščemo tisto linearno kombinacijo vektorjev, pri kateri bo srednja kvadratna napaka preslikave poljubnega vektorja v izbran podprostor najmanjša. Linearno kombinacijo vektorjev, ki tvorijo podprostor imenujemo lastne slike. Lastne slike so v bistvu lastni vektorji kovariančne matrike, izračunane iz vektorsko predstavljenih slik učne množice.

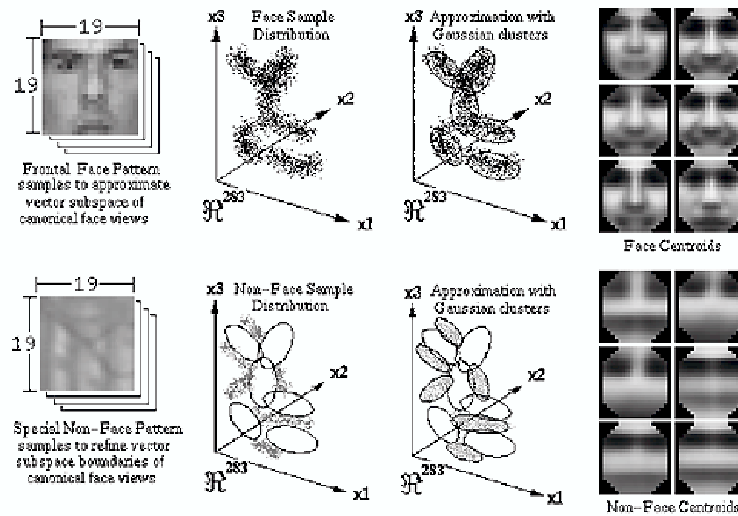
Poizkusi z učno množico 100 slik so pokazali, da lahko sliko velikosti 91×50 slikovnih elementov učinkovito predstavimo s pomočjo 50 lastnih slik, pri čemer ohranimo dovolj podobnosti na slikah. V tem poskusu se je ohranilo 95% variance učne množice.

Turk in Pentland [70] sta analizo glavnih komponent uporabila tako za iskanje, kot tudi za prepoznavo obrazov. Iz učne množice sta izračunala lastne slike, imenovane tudi lastni obrazi, ki so predstavljale podprostor vseh slik, oziroma tako imenovan obrazni prostor. Slike obrazov, ki sta jih predstavila v obraznem prostoru, sta združila v skupine. Podobno sta tudi slike neobrazov predstavila v istem obraznem prostoru in jih združila. Medtem, ko se slike obrazov, predstavljene v obraznem prostoru, niso veliko spremenile, so se slike neobrazov spremenile. Ali slika vsebuje obraze, sta tako ugotovila s pomočjo izračuna razdalje med sliko in obraznim prostorom. Lokalni minimumi na tako izračunani matriki razdalj so predstavljali obraze. Lastne vektorje za razčlenitev in povezljivost se še vedno pogosto uporabljala na področju iskanja in prepoznave obrazov ter za iskanje značilk obraza.

2.4.2 Metode na osnovi porazdelitev

Sung in Poggio [66] sta razvila metodo za iskanje obrazov, ki temelji na sistemu porazdelitvenih modelov. Pokazala sta, kako se lahko naučimo porazdeljenosti slikovnega vzorca enega razreda s pomočjo negativnih in pozitivnih primerov tega razreda. Njun sistem je sestavljen iz dveh komponent: iz porazdelitvenih modelov za razred obrazov oziroma za razred neobrazov (Slika 2.3) ter klasifikatorja na osnovi večnivojskega perceptrona.

Posamezno učno sliko obraza in neobraza sta najprej normalizirala in zmanjšala na velikost 19×19 slikovnih elementov. Sliko sta nato obravnavala kot vektor dolžine 361 elementov. Tako zgrajene vektorje, oziroma vzorce obrazov in neobrazov, sta na podlagi *k-means* algoritma razdelila v šest skupin. Vsako skupino sta predstavila s pomočjo večdimenzijske Gaussove funkcije s srednjo vrednostjo in kovariančno matriko. Na podlagi tako zgrajenega porazdelitvenega modela vzorcev obraza in porazdelitvenega modela vzorcev neobraza sta določila dve matriki razdalj. V prvi matriki so Mahalanobisove razdalje med testnim vzorcem in središčem posamezne skupine iz modela. Razdalje so merjene v dimenzijsko zmanjšanem podprostoru, ki ga določa 75 največjih lastnih vektor-



Slika 2.3: Porazdelitvena modela za razred obrazov in neobrazov, ki sta ju uporabila Sung in Poggio v svojem delu [66]. Posamezen model sestavlja mešanica Gaussovih funkcij. Na desni strani so prikazani centri Gaussovih funkcij.

jev modela. Druga matrika razdalj pa vsebuje evklidske razdalje med testnim vzorcem in njeno projekcijo v omenjeni 75 dimenzijski podprostor. 12 tako določenih parov razdalj nato pripeljeta na vhod nevronske mreže, ki ločuje obraze od neobrazov. Klasifikator je naučen s standardnim postopkom vzratnega razširjanja (ang. backpropagation) na 47316 učnih primerih. Od tega je bilo 4150 pozitivnih primerov, ki so predstavljali obraz, ostali primeri pa niso vsebovali obraza. Poudariti je potrebno, da pozitivnih učnih primerov, ki predstavljajo obraze ni težko izbrati. Težje je izbrati reprezentativne primere, ki ne vsebujejo obrazov. Zaradi tega je bila predlagana metoda razmnoževanja učnih primerov (ang. bootstrap), ki dopolnjuje učno množico neobrazov med samim učenjem. S pomočjo nevronske mreže, naučene na podlagi majhne učne množice neobrazov, izvedemo iskanje nad poljubnim zaporedjem slik in izberemo napačno klasificirane primere. Te primere nato dodamo v učno množico kot nove primere neobrazov. Takšna tehnika razmnoževanja učnih primerov rešuje problem slabo določenih učnih primerov neobrazov in je bila kasneje še velikokrat uporabljena [48, 53].

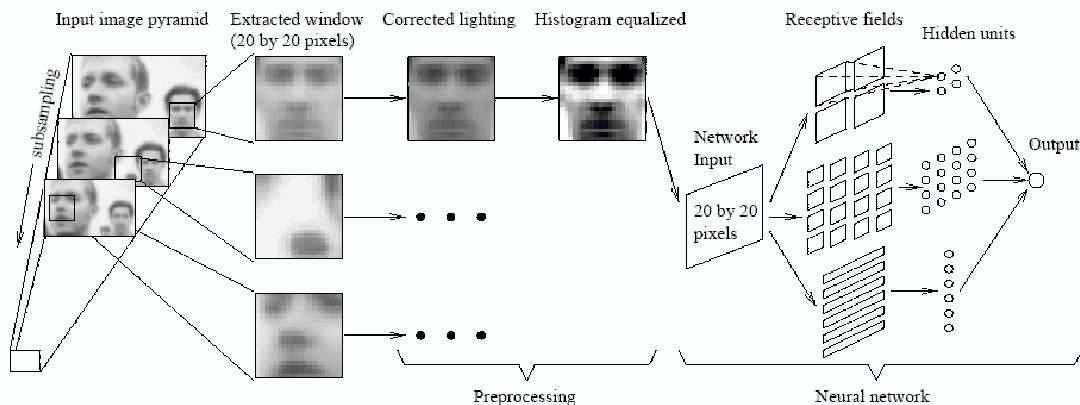
2.4.3 Nevronske mreže

Nevronske mreže se velikokrat uporabljajo pri problemih prepoznavanja vzorcev, kot so na primer optično prepoznavanje pisave, prepoznavanje objektov ali avtomatsko krmiljenje robotov. Problem iskanja obrazov uvrščamo med probleme prepoznavanja vzorcev z dvema razredoma: razredom obrazov in razredom neobrazov, ki vsebuje primere, na katerih ni obraza. Za tako specializiran problem je bilo predlaganih veliko različnih arhitektur nevronskih mrež. Prednost uporabe nevronskih mrež pri učenju obraznega vzorca je vsekakor sposobnost mreže, da se lahko nauči kompleksnosti vzorca. Za doseganje sprejemljive uspešnosti učenja mreže, pa moramo znati zgraditi ustrezno mrežo: določiti moramo pravo število nivojev in število nevronov na vsakem nivoju, primerno stopnjo učenja, itd.

Med vsemi metodami za iskanje obrazov, ki temeljijo na nevronskih mrežah, je najpomembnejše delo avtorja Rowleya s sodelavci [53, 54, 55]. S pomočjo večnivojske nevronske mreže se je naučil razlikovati med obrazi in neobrazi, medtem ko je Sung [65] uporabil nevronske mreže za določitev diskriminantne funkcije, ki ločuje med obrazi in neobrazi.

Za razliko od drugih avtorjev [8, 71], ki so za učenje uporabili eno ali dve nevronske mreži, je Rowley učinkovitost povečal z uporabo več nevronskih mrež in dodatnimi metodami, s katerimi je izbiral med posameznimi izhodi nevronskih mrež. Njegovo metodo sestavljata dve glavni komponenti: nevronske mreže za iskanje obrazov in odločitveni model, ki določi končno odločitev med vsemi izhodi nevronskih mrež.

Na sliki 2.4 lahko vidimo primer nevronske mreže, ki na vhod sprejme sliko velikosti 20×20 , na izhodu pa vrača rezultat v intervalu od -1 do 1. Torej nam za dano sliko na vhodu vrne nevronska mreža vrednost blizu -1, če na vhodu ni obraza, ali pa vrednost blizu 1, če je na vhodu slika obraza. S pomočjo tako definiranega klasifikatorja lahko obraze na sliki najdemo tako, da preiščemo posamezne dele celotne slike. Če hočemo poiskati tudi obraze, ki so večji od velikosti 20×20 , pa moramo vhodno sliko primerno zmanjšati, za iskanje različno obrnjenih obrazov pa moramo preiskovati ustrezno obrnjeno sliko. Za učenje nevronske mreže je Rowley uporabil okrog 1050 učnih primerov slik obrazov različnih velikosti, različnih položajev in različnih osvetlitev. Na podlagi ročno označenih značilk obrazov je slike ustrezno normaliziral in svetlobno popravil. V drugem delu svoje metode je poiskal obraze, ki so bili večkrat odkriti, in se odločil med posameznimi izhodi



Slika 2.4: Diagram nevronske mreže, ki jo je uporabil Rowley [53] v svoji metodi za iskanje obrazov na sliki.

različnih nevronskih mrež. Za odločevanje je uporabil preprosto odločitveno shemo, ki je temeljila na logičnih operacijah AND in OR in tehniki glasovanja (ang. voting).

Sistem za iskanje obrazov, ki ga je razvil Rowley s sodelavci, je računsko manj zahteven od sistema, ki sta ga razvila Sung in Poggio [66] in dosega boljše rezultate. Oba sistema pa imata to pomanjkljivost, da lahko najdeta samo obraze, ki so frontalni in pokončni. Rowley je kasneje svojo metodo nadgradil tako, da najde tudi obraze obrnjene za določen kot. To je dosegel s pomočjo nevronske mreže, ki ugotovi, za koliko je vhodna slika obrnjena. Pravilno obrnjeno sliko nato pripelje na vhod prej predstavljene nevronske mreže, ki razlikuje med obrazi in neobrazi. Takšen sistem je od prejšnjega slabši pri pokončnih obrazih, najde pa 76.9% vseh različno obrnjenih obrazov z zelo majhno napako.

2.4.4 Metoda podpornih vektorjev

Metodo podpornih vektorjev (ang. Support Vector Machines, SVM) je na obrazih prvič uporabil Osuna [48].

Metoda podpornih vektorjev omogoča nov način učenja, drugačen kot smo ga do sedaj poznali pri klasifikatorjih, kot so Bayesov ali nevronske mreže. Večina postopkov učenja temelji namreč na minimizaciji klasifikacijske napake (t.i. empiričnega tveg-

anja), metoda podpornih vektorjev pa deluje na drugačen način. Njen cilj je minimizirati zgornjo mejo pričakovane napake posplošitve, kar imenujemo minimizacija strukturnega tveganja. Klasifikator na podlagi SVM-ja je linearni klasifikator, ki poskuša najti takšno hiperravnino, da je pričakovana klasifikacijska napaka, za nek neznan primer, najmanjša. Takšno optimalno hiperravnino določimo s pomočjo utežene kombinacije majhne domene učnih vektorjev, ki jih imenujemo tudi podporni vektorji. Na iskanje optimalne hiperravnine lahko gledamo tudi kot na reševanje linearnega kvadratičnega optimizacijskega problema z omejitvami.

Metoda podpornih vektorjev je tako časovno kot spominsko zelo zahtevna, kljub temu pa je Osuna [48] na njeni osnovi izdelal učinkovito metodo za iskanje obrazov. Metodo je učil na podlagi 10 milijonov učnih primerov velikosti 19×19 in dobil zelo majhno število napak pri iskanju. Njegova metoda je tudi 30 krat hitrejša od metode avtorjev Sung in Poggio [66]. Kasneje so metodo podpornih vektorjev v kombinaciji z metodo valčkov (ang. wavelets) večkrat uporabili za iskanje obrazov in pešcev na slikah.

2.4.5 Naivni Bayesov klasifikator

V nasprotju s prej opisanimi metodami, ki modelirajo globalno predstavitev celotnega obraza, sta Schneiderman in Kanade [59] s pomočjo naivnega Bayesovega klasifikatorja ocenila skupno verjetnost lokalnih značilk (t.i. podregij obraza) in njihov položaj pri različnih velikostih obraza. Upoštevala sta lokalni videz objekta, saj so nekateri lokalni vzorci bolj pomembni kot stali. Na primer pri obrazu je vzorec okrog oči veliko bolj izrazit od vzorca na licih. Naivni Bayesov klasifikator, ki predpostavlja neodvisnost vhodnih vzorcev, sta uporabila iz dveh razlogov. Prvič, ker dobro določi samo funkcijo porazdelitve za posamezne podregije. Drugič pa zato, ker lahko s pomočjo naivnega Bayesovega klasifikatorja določimo posteriorno verjetnost (ang. posterior probability). Pri različnih velikostih sliko najprej razdelita na podregije. Te podregije nato s pomočjo metode PCA preslikata v nižjedimezijski prostor in jih združita v končno množico vzorcev. Vsaki tako preslikani podregiji priredita verjetnost na podlagi preslikanih vzorcev. Metoda določi, da je na sliki obraz, če je razmerje verjetnosti večje od razmerja apriornih verjetnosti. S 93% točnostjo se metoda lahko primerja z metodo na osnovi nevronske mreže, s tem, da metoda najde tudi delno nagnjene obraze in obraze s profila.

2.5 Zaključek

S pomočjo predlaganih tehnik, ki smo jih predstavili v tem poglavju, je bilo razvitih že veliko metod za iskanje obrazov na sliki. Primerjava njihove učinkovitosti pa ni enostavna, ker veliko metod za preizkus svoje učinkovitosti uporablja različne testne množice. Pri tistih, ki uporabljajo iste množice slik pa se pojavita dva druga problema. Kot prvič, raziskovalci različno intepetirajo, kaj je uspešno najden obraz, in kot drugič, za učenje uporabljajo različne učne množice.

Za preizkus metod sta največkrat uporabljeni testna množica MIT, avtorjev Sung in Poggio [66] in testna množica CMU, avtorjev Rowley s sodelavci [55]. Množica MIT vsebuje 125 slik z 483 obrazi in množica CMU vsebuje 32 slik z 136 obrazi. Tudi za preizkus metode BDF (poglavje 5) smo uporabili testne slike, ki so vsebovale samo frontalne obraze iz teh dveh testnih množic.

V tabeli 2.1 so združeni rezultati pomembnejših metod na osnovi videza, kot so metoda na osnovi porazdelitev [66], metoda Nevronskih mrež [55], metoda podpornih vektorjev [48] in Naivni Bayesov klasifikator, avtorjev Schneiderman in Kanade [59]. Vse metode so bile preizkušene vsaj na eni izmed prej omenjenih testnih množicah MIT in CMU.

Metoda	testna baza MIT		testna baza CMU	
	uspešnost(TP)	napačni(FP)	uspešnost(TP)	napačni(FP)
na osnovi porazdelitev	ni podatka	ni podatka	81.9%	13
nevronske mreže	92.5%	862	90.3%	42
podporni vektorji	ni podatka	ni podatka	74.2%	20
naivni Bayes	93.0%	88	91.2%	12

Tabela 2.1: Preizkus različnih metod za iskanje obrazov na testni množici MIT in testni množici CMU.

Poglavje 3

Različne metode in barvni prostori za segmentacijo kožne barve

Ko gradimo sistem za iskanje obrazov na osnovi kožne barve se običajno soočimo s tremi glavnimi problemi. Najprej, kakšen barvni prostor naj izberemo, nato, kako naj porazdelitev barve kože modeliramo, in končno, na kakšen način naj nadalje obdelamo kožne regije na sliki.

V poglavju 3.1 bom opisal posamezne barvne prostore s poudarkom na tistih, ki jih lahko uporabimo za segmentacijo kožne barve. V naslednjem poglavju 3.2 pa bom naredil pregled različnih metod, s pomočjo katerih lahko ločimo kožno barvo od drugih barv.

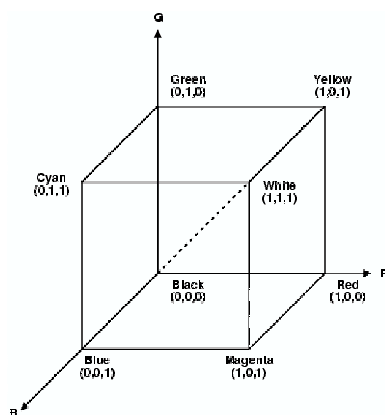
3.1 Pregled barvnih prostorov

Preučevanje barv, razvoj računalniške grafike in različni standardi za prenos video signalov so prinesli različne barvne prostore z različnimi lastnostmi. Veliko barvnih prostorov se uspešno uporablja tudi pri modeliranju kože barve za potrebe iskanja obrazov na sliki. V tem poglavju bomo preučili različne barve prostore in njihove lastnosti, ki ugodno vplivajo na iskanje kožne barve obraza na sliki. V zaključku bomo predlagali najboljši in najučinkovitejši barvni prostor, ki ga bomo v nadaljevanju uporabili pri implementaciji različnih metod.

3.1.1 Barvni prostor RGB (ang. Red, Green, Blue)

Kratice označujejo posamezne barvne kanale: rdeči, zeleni in modri barvni kanal. Uporablja se predvsem v sistemih, ki za opisovanje barv uporabljajo katodno cev (ang. Cathode Ray Tube, CRT). Takšne naprave so predvsem analogni računalniški zasloni. Barvni prostor RGB uvrščamo med barvne prostore, ki so odvisni od naprave. Pri zaslonih to pomeni, da je prikazana barva odvisna od električnih karakteristik katodne cevi in samih nastavitev zaslona. Barvo opišemo s pomočjo kombinacije treh barvnih žarkov (rdeč, zelen, moder), kjer predstavlja posamezna vrednost količino posamezne barve v odstotkih. Vrednosti za posamezni kanal so tako v intervalu $[0,1]$. V računalništvu je posamezna vrednost barve predstavljena kot 8 bitna vrednost ($2^8 = 256$ različnih vrednosti), zato posamezno barvo predstavimo v računalništvu tudi z vrednostmi v intervalu $[0, 255]$, kjer 100% zastopanost posamezne barve predstavlja vrednost 255.

Barvni prostor RGB je eden izmed najbolj razširjenih barvnih prostorov za obdelavo in hranjenje digitalnih slik. Barvni prostor RGB si lahko predstavimo s pomočjo kocke v kartezijskem koordinatnem sistemu, kjer os x običajno predstavlja rdečo komponento (R), os y modro (B) in os z zeleno (G) (slika 3.1). Diagonala ki povezuje črno barvo (ang. black), v točki $(0,0,0)$ in belo (ang. white) v točki $(1,1,1)$, opisuje vse možne sivine (črtkana črta).



Slika 3.1: Predstavitev barvnega prostora RGB s pomočjo kocke.

Barvni prostor RGB zaradi svojih lastnosti ni primeren za barvne analize ter za algo-

ritme za razpoznavo predmetov. Neugodne lastnosti barvnega prostora RGB so: velika korelacija med posameznimi barvnimi kanali, značilna neenotna občutljivosti in mešanje kromatičnosti barve (ang. chrominance) s svetlostjo (ang. luminance). Posamezne vrednosti barvnih kanalov, R, G in B namreč ne predstavljajo samo barve ampak tudi svetlost. Če sta dve barvi (r_1, g_1, b_1) in (r_2, g_2, b_2) med seboj v istem razmerju:

$$\frac{r_1}{r_2} = \frac{g_1}{g_2} = \frac{b_1}{b_2}$$

označujeta isto barvo, imata pa lahko različno svetlost.

Človeški sistem zaznavanja barv se prilagaja na različno svetlost in na različne vire osvetlitev, tako da je zaznavanje barve v različno osvetljenih okoljih vedno konstantno. Iz kožne barve lahko na takšen način izločimo osvetlitev tudi v primeru, ko sprejemamo manj barvne informacije. Za razlikovanje kožne barve je torej potrebna samo barva, brez informacije svetlosti. Zaradi tega so se, pri sistemih za segmentacijo barve, uveljavili predvsem barvni prostori, ki vsebujejo samo kromatičnost barve. Zaradi svoje enostavnosti pa se velikokrat uporablja tudi barvni prostor RGB [5, 28]. Uporabljen je tudi za segmentacijo kožne barve v instalaciji 15 sekund slave [63].

3.1.2 Normaliziran barvni prostor RGB ali barvni prostor rg

Normaliziran barvni prostor RGB lahko zelo enostavno dobimo s pomočjo normalizacije treh barvnih vrednosti barvnega prostora RGB:

$$r = \frac{R}{R + G + B}, \quad g = \frac{G}{R + G + B}, \quad b = \frac{B}{R + G + B}. \quad (3.1)$$

Ker je vsota vseh treh komponent znana ($r + g + b = 1$), je tretja komponenta odveč in jo lahko izpustimo. S tem zmanjšamo dimenzijo barvnega prostora iz 3D v 2D. Preostali komponenti: r in g , običajno imenujemo "čisti barvi" (ang. pure colors).

Z normalizacijo zmanjšamo tudi odvisnost tako predstavljene barve od svetlosti (ang. brightness). Pri predmetih brez leska in pri neupoštevanju okoliške svetlobe je barva, predstavljena v tem barvnim prostoru, nespremenljiva na orientacijo opazovanega predmeta glede na vir svetlobe. Takšna lastnost je pri iskanju objektov na slikah pri različni osvetlitvi zelo dobrodošla. Če poudarimo še enostavnost preslikave barvnega prostora

rg iz barvnega prostora RGB vidimo, zakaj je ta barvni prostor tako priljubljen in tudi največkrat uporabljen za segmentacijo kožne barve na sliki [7, 84, 47, 77].

3.1.3 HSI, HSV, HSL - Hue, Saturation, Intensity (Value, Lightness)

Posamezne komponente barvnih prostorov HSI , V , L določajo posamezne lastnosti barve, ki se uporabljajo tudi v slikarstvu. Potreba po teh barvnih prostorih se je pojavila z željo, da bi lahko posamezno lastnost barve določili številčno. Prva vrednost predstavlja barvni ton (ang. hue). Barvni ton je človeški občutek s katerim presojaмо ali je neka površina podobna eni barvi oziroma kombinaciji dveh barv, ki jih lahko zaznavamo (rdeči, zeleni, modri). Je funkcija valovne dolžine in običajno barvni ton imenujemo kar barva (rdeča, zelena, rumena, modra, itd.). Druga vrednost predstavlja barvno nasičenost (ang. saturation). Barvno nasičenost lahko opredelimo kot barvitost površine glede na njeno svetlost ali kot intenzivnost opazovane barve. Tretja vrednost (I , L , V) predstavlja svetlost barve (ang. lightness). Ker je ta vrednost povezana z različnimi pojmi, kot so intensity, brightness, luminance in lightness, si v nadaljevanju pogledjmo njihov pomen.

Intensity (svetilnost): Svetlobni izvor izžareva svetlobni tok v različnih smereh ter z različno jakostjo. Vidno intenziteto radiacije v izbrani smeri imenujemo svetilnost (ang. intensity). Pri slikah merimo s svetilnostjo moč na določenem intervalu elektromagnetnega spektra in običajno nas zanima tista moč, ki prihaja iz ali vpada na določeno površino. Svetilnost običajno imenujemo merilo za linearno svetlobo, ki jo predstavimo z enoto: vati na kvadratni meter.

Brightness (svetlost): Ta izraz je bil definiran s strani CIE (fr. Commission Internationale de L'Eclairage), kot občutek s katerim človek zazna, da je neka površina bolj ali manj svetla. Ta vrednost občutka je subjektivne narave in nima objektivne merilne enote.

Luminance (svetlost): CIE je definiral "luminance", kot moč svetilnosti, utežene s funkcijo spektralne občutljivosti značilne za človeški vid. Svetlost je v sorazmerju s fizično močjo. V tem pogledu je enaka svetilnosti (ang. intensity). Ampak spektralna sestava svetilnosti je v tem primeru poveza z občutljivostjo človeškega vida na svetlost. Svetlost

lahko izračunamo kot vsoto primerno utežene linearne svetlobe rdeče, zelene in modre barvne komponente (R, G, B). Po definiciji je svetlost definirana kot:

$$Y = 0.2125R + 0.7154G + 0.0721B.$$

Označujemo jo tudi s simbolom Y_{709} , kjer številka 709 označuje standard, po katerem so določene posamezne uteži.

Lightness (svetlost): S človeškim vidom zaznavamo svetlobo nelinearno. Na primer, izvor svetlobe, ki ima svetilnost samo 18% manjšo od referenčne vrednosti svetilnosti, dojamemo kot pol svetlejši. Svetlost (ang. lightness), ki predstavlja dojemljivost človeškega vida za svetilnost (Y), je CIE definiral kot:

$$L^* = 116 * \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} - 16, \quad 0.008856 < \frac{Y}{Y_n},$$

kjer vrednost Y_n označuje svetilnost referenčne bele barve. Vrednosti $Y/Y_n > 0.008856$ pripadajo zelo majhnim vrednostim Y in označujejo zelo temen spekter. V tem primeru je svetlost kar linearna s koeficientom 903.3. Za praktično uporabo majhne vrednosti svetilnosti niso pomembne, je pa važno da jo omejimo na vrednost 0. Vrednosti, ki jih lahko zavzame svetlost so v intervalu $[0,1]$.

Kot smo povedali že v uvodu, sestavljajo barvne prostore HSI, V, L tri komponente, ki označujejo barvni ton, nasičenost in svetlost barve.

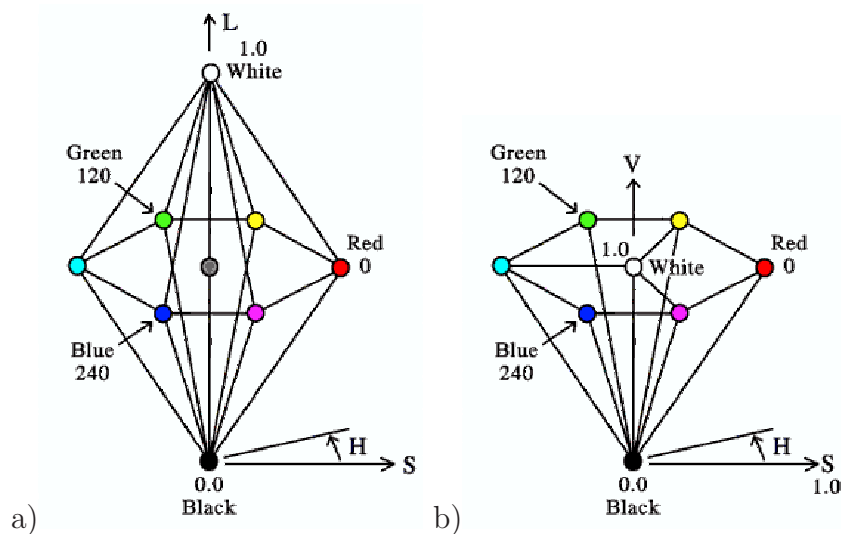
Barvni ton (ang. hue) lahko zavzame vrednosti v intervalu $[0,1]$. V tem intervalu se zvrstijo vse barve od rdeče, rumene, zelene, modre, vijoličaste in nazaj do rdeče. Rdečo barvo označujeta namreč dve vrednosti 0 in 1.

Tudi barvna nasičenost (ang. saturation) zavzame vrednosti v intervalu $[0,1]$. Izbrana barva v določenem barvnem tonu je lahko zelo nasičena (ne vsebuje bele komponente) ali pa nenasičena (vsebuje odtenke sive barve). Za barvno nasičenost lahko rečemo tudi, da označuje čistost oziroma živost barve. Majhne vrednosti označujejo več bele barve, kar predstavlja bolj pastelne barve. Zelo majhne vrednosti pa predstavljajo "izprane" barve. Za čiste barvne tone mora biti vrednost nasičenosti maksimalna. Tako imenovane akromatične barve (črna, bela, siva) imajo nasičenost enako 0.

Svetlost (ang. lightness) lahko prav tako zavzame vrednosti v intervalu $[0,1]$ in predstavlja delež črne barve v barvnem tonu oziroma kontrolira svetlost barve. Barva z veliko vrednostjo je bolj svetla, z manjšo pa manj. Maksimalna vrednost vedno predstavlja belo barvo ne glede na ton barve, minimalna vrednost pa vedno črno barvo.

Barvni prostor HSL lahko prikažemo z modelom dvojnega stožca (Slika 3.2a). Vertikalna os predstavlja svetlost L (ang. lightness). Začne se na dnu, pri vrednosti 0, ki označuje črno barvo in konča na vrhu pri vrednosti 1, ki označuje belo barvo. Oddaljenost od središča osi predstavlja nasičenost S (ang. saturation). Vse točke na vertikalni osi imajo nasičenost enako 0, torej predstavljajo vse odtenke sivine. Točke, ki so bolj oddaljene od osi imajo večjo nasičenost in predstavljajo bolj "čiste" barve. Parameter H označuje barvni ton, meri pa se kot velikost kota pri rotaciji okrog šesterokotnika.

Barvni prostor HSV pa lahko prikažemo z modelom enojnega stožca (slika 3.2b). Vrednost V ravno tako predstavlja svetlost L (ang. lightness) in zavzema vrednosti 0, na dnu modela, ki predstavlja črno barvo in vrednost 1 na vrhni ploskvi, ki predstavlja belo barvo.



Slika 3.2: Predstavitev barvnih prostorov HSL in HSV.

Zaradi intuitivnosti posameznih komponent in eksplicitnega razlikovanja informacije o svetlosti (ang. luminance) od informacije o barvi oziroma kromatočnosti (ang. chromi-

nance), so ti barvni prostori zelo uporabni za segmenacijo barve [84]. Omenjeni barvni prostori imajo še druge dobre lastnosti, kot so n.pr. neodvisnost na premočno belo svetlobo in neodvisnost na prostorsko svetlobo za površine brez leska in površine, ki so neposredno obrnjene na izvor svetlobe.

Znane pa so tudi nekatere nezaželenne lastnosti [52]. Te vključujejo nepovezanost barvnega tona, ki je določen s kotnimi stopinjami in neskladnost izračunane svetlosti barve s človeškim zaznavanjem svetlosti. Na primer, enako svetli barvi rumena in modra se po izračunu med seboj razlikujeta za faktor šest.

Posamezne komponente barvnega prostora HSV dobimo s pomočjo transformacije:

$$\begin{aligned} H &= \arccos \frac{\frac{1}{2}((R - G) + (R - B))}{\sqrt{(R - G)^2 + (R - G)(G - B)}}, \\ S &= 1 - 3 \frac{\min(R, G, B)}{R + G + B}, \\ V &= \frac{1}{3}(R + G + B). \end{aligned}$$

Alternativni izračun za komponenti H in S vključuje v izračun tudi logaritem, kar naj bi zmanjšalo odvisnost kromatičnosti barve od njene svetlosti.

Polarne koordinate, v katerih merimo vrednost H , naredijo te barvne prostore nepri-merne za modeliranje barvne kože na podlagi učne množice, zaradi tega se uporablja predstavitev teh barvnih prostorov v kartezijskih koordinatah: $X = S * \cos H$, $Y = S * \sin H$.

3.1.4 TSL - Tint, Saturation, Lightness

Normaliziran barvni prostor TSL je spremenjen normaliziran barvni prostor RGB, tako da so vrednosti bolj intuitivne glede na izbiro barvnega tona in nasičenosti. Pri temu prostoru imamo opravka z RGB komponentami, ki so gama-korekcijsko popravljene in tako popravljene vrednosti bomo označevali z R' , G' , B' . Za transformacijo iz barvnega prostora $R'G'B'$ v barvni prostor TSL uporabimo sledeče formule:

$$\begin{aligned}
S &= \sqrt{\frac{9(r'^2 + g'^2)}{5}}, \\
T &= \begin{cases} \arctan(r'/g')/2\pi + 1/4; & g' > 0, \\ \arctan(r'/g')/2\pi + 3/4; & g' < 0, \\ 0; & g' = 0 \end{cases} \\
L &= 0.299R' + 0.587G' + 0.114B',
\end{aligned}$$

kjer je $r' = r - 1/3$, $g' = g - 1/3$. Vrednosti r in g pa izhajata iz enačbe 3.1. Vrednost T predstavlja barvni ton, vrednost S nasičenost barve in vrednost L svetlost barve (za gama-korekcijsko popravljene barvne vrednosti R, G, B). Vrednosti S , T in L pred uporabo normaliziramo na interval $[0,1]$.

Terrillon [69] je primerjal med seboj devet različnih barvnih prostorov. Njihove barvne komponente, ki predstavljajo kromatičnost, je uporabil za modeliranje barve kože s pomočjo Gaussove porazdelitvene funkcije. Ugotovil je, da je normaliziran barvni prostor TSL daleč najboljši za takšno opravilo. Vrednosti T in S namreč za barvo kože tvorijo dovolj ozek in enoten prostor, da ga lahko opišemo z enim samim Gaussovim modelom porazdelitve. Tudi Brown [7] je pri svojem delu uporabil ta barvni prostor.

3.1.5 Y'CrCb barvni prostor

V to skupino barvnih prostorov sodijo še podobni barvni prostori kot so $Y'U'V'$ ali EBU , $Y'I'Q'$, itd. Uporabljajo se pri televizijskih oddajnih sistemih. $Y'I'Q'$ in $Y'U'V'$ (EBU) barvna prostora sta analogna in se uporabljata za televizijska standarda NTSC in PAL, medtem ko je $Y'CrCb$ digitalni standard. Njihova glavna značilnost je, da ločijo RGB komponente na svetlost (Y') in barvo (Cr, Cb). To je lahko uporabno v aplikacijah za stiskanje slik (tako digitalnih kot analognih). Ti prostori pa niso pretirano intuitivni za izbiro barv. Tudi pri teh prostorih uporabljamo gama-korekcijsko popravljene komponente barvnega prostora RGB, ki jih lahko spremenimo v sami video kameri ali s pomočjo korekcijske tabele. Vsaka od teh komponent prispeva drugačen delež k osvetlitvi, ki jo

oko zaznava, zaradi treh različnih tipov čepkov v človeškem očesu. Novo vrednost Y' , ki se pojavlja v teh barvnih prostorih tako izračunamo kot:

$$Y' = 0.299R' + 0.587G' + 0.114B'.$$

Vrednost Y' imenujemo "Luma" in označuje svetlost barve. Ker so vse tri popravljene RGB vrednosti normalizirane je tudi vrednost Y' normalizirana na območju $[0,1]$. Komponenti, ki označujeta barvo oziroma kromatičnost barve pa označimo z U' in V' . Izračunamo pa ju kot razliko: $U' = B' - Y'$ in $V' = R' - Y'$. Njunjo vrednost si lahko predstavljamo kot vsebnost ali pomanjkanje modre oziroma rdeče komponente v barvi. Ker so vse vrednosti R' , G' , B' in Y' normalizirane na območju $[0,1]$ ni težko ugotoviti, da je vrednost U' določena na območju $[-0.886, +0.886]$ in vrednost V' na območju $[-0.701, +0.701]$. Normalizirani vrednosti za U' in V' označujemo s komponentama Cb in Cr :

$$Cb = (U'/2) + 0.5, \quad Cr = (V'/1.6) + 0.5.$$

Omenjeni barvni format $Y'CbCr$ se običajno uporablja za digitalni video in računalniško grafiko. Metode za stiskanje slik in videa predpostavljajo, da so slike predstavljene v tem barvnem prostoru. Posebnost tega barvnega prostora je v tem, da je sprememba svetlosti (Y') veliko bolj opazna kot sprememba barve (Cb, Cr). Metode za stiskanje zato poskušajo pri zmanjševanju informacij čimbolj ohraniti komponento Y , medtem ko lahko informacijo o barvi zmanjšajo brez večjih posledic za kvaliteto. Zaradi enostavne pretvorbe RGB signala in razlikovanja svetlosti od barve je tudi $Y'CbCr$ barvni prostor zelo priljubljen pri modeliranju kožne barve [50, 84, 45, 25, 1, 10].

Med ostalimi barvnimi prostori, ki so tudi linearne preslikave barvnega prostora RGB in bi jih zaradi tega lahko uvrščali v to skupino, so bili za segmentiranje kožne barve uporabljeni še barvni prostori YES [56], YUV [42] in YIQ [5, 17].

3.1.6 Barvna prostora CIE-Luv in CIE-Lab

Izraz "kožna barva" ne označuje fizične lastnosti objekta, ampak je bolj predmet zaznavanja in zaradi tega je ta pojem bolj povezan s subjektivnostjo človeškega zaznavanja.

Zaradi tega bi predstavitev barve, ki je bližje človeškemu zaznavanju, lahko pripomogla k boljšim metodam za iskanje kožne barve. Barvna prostora CIE Lab in CIE Luv sta skoraj linearna glede na človeško zaznavanje barv oziroma se temu od vseh barvnih prostorov še najbolj približata. Izhajata iz barvnega sistema, ki ga je definiral CIE (fr. Commission Internationale de L'Eclairage).

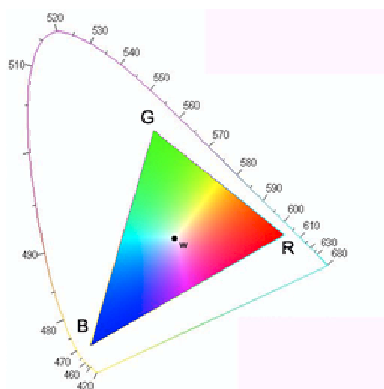
Barvni sistem XYZ

Že leta 1931 je CIE standardiziral barvni sistem XYZ . Definiral ga je tako, da klasificira barvo na osnovi sistema človeškega vida. S primerjanjem barve spektra s tremi barvnimi svetlobnimi viri je CIE izmeril občutljivost treh pasovnih širin v očesu. Porazdelitev moči spektra barve je razbita s funkcijami občutljivosti tako, da dobimo kot rezultat tri vrednosti X , Y in Z . Te tri vrednosti ne predstavljajo nobene prave barve, ampak lahko z njihovo linearno kombinacijo, $xX + yY + zZ$ dobimo katero koli barvo. Vrednosti x, y, z so pozitivne in velja $x + y + z = 1$. Vrednost Y predstavlja svetlost, ostali dve komponenti X in Z pa opisujeta barvo oziroma kromatičnost. Posamezno barvo v tem sistemu lahko določimo s pomočjo vrednosti x in y in tako dobimo barvni prostor Yxy :

$$x = \frac{X}{X + Y + Z}, \quad y = \frac{Y}{X + Y + Z}.$$

Določimo lahko tudi tretjo koordinato, vendar je ta nepotrebna, saj velja $x + y + z = 1$. Vse barve lahko tako prikažemo na pripadajoči ravnini na tako imenovanem CIE kromatičnem diagramu (Slika 3.3).

Čiste barve ležijo na krivulji, ki označuje valovne dolžine od 420nm do 660nm. Krivulja ima obliko konjskega kopita. Označena točka \mathbf{w} na sredini označuje čisto belo barvo, ki ni nasičena (vrednost saturation je 0) in se v praksi uporablja za primerjavo med belimi barvami. Pripadajoči trikotnik, ki leži znotraj krivulje, pripada RGB barvni lestvici. Robovi trikotnika pripadajo posameznim barvnim komponentam R , G in B . Posamezne barve pa dobimo z linearno kombinacijo posameznih komponent, kar predstavlja notranjost trikotnika. Iz slike lahko vidimo, da katere koli tri vrednosti komponent barvnega prostora RGB ne morejo zajeti vseh barv, ki jih predstavlja notranjost krivulje. Iz tega sledi, da s pomočjo barvnega prostora RGB ne moremo predstaviti vseh barv. Težava Yxy



Slika 3.3: CIE kromatični diagram za prikaz barv v Yxy barvnem prostoru. Trikotnik označuje barvni prostor RGB.

barvnega prostora pa je močna nelinearnost glede velikosti enote, ki predstavlja razliko med dvema barvama. Razlika med posameznima barvama namreč ni konstantna.

CIE-YUV

CIE- Yuv (1960) je linearna transformacija prostora Yxy z namenom zmanjšati nelinearnost glede velikosti enote. S tem sistemom nelinearnost zmanjšamo, a še vedno ne dovolj, da bi bila razdalja med dvema barvama konstantna. Transformacija je sledeča:

$$u = \frac{2x}{6y - x + 1.5}, \quad v = \frac{3y}{6y - x + 1.5}.$$

Določili bi lahko tudi tretjo koordinato, a je to nepotrebno, saj velja: $u + v + w = 1$.

Da bi obstoječo nelinearnost zmanjšali se je pojavila še linearna transformacija $Yu'v'$, ki pa je še vedno relativno neučinkovita z vidika zmanjšanja nelinearnosti. Vrednosti Y in u sta ostali nespremenjeni:

$$u' = u = \frac{2x}{6y - x + 1.5}, \quad v' = 1.5v \frac{4.5y}{6y - x + 1.5}.$$

CIE-Luv

Osnova za barvni prostor CIE-*Luv* je bil barvni prostor CIE- $Y u' v'$. Glavni namen uvedbe tega barvnega prostora je bil še vedno linearizacija vektorja razlik med barvami. CIE-*Luv* je nelinearen barvni prostor, ki ima tudi obratno transformacijo. Barvna informacija je centralizirana na vrednost bele točke barvnega sistema, ki je v enačbi podpisana s črko n . Nelinearna transformacija vrednosti (Y/Y_n) , ki označuje relativno svetlost poskuša posnemati logaritemsko odzivnost očesa na svetlobo:

$$L = \begin{cases} 116 \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} - 16; & \text{ce } \frac{Y}{Y_n} > 0.008856 \\ 903.3 \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} - 16; & \text{ce } \frac{Y}{Y_n} \leq 0.008856 \end{cases}$$

$$u = 13(L)(u' - u'_n), \quad v = 13(L)(v' - v'_n).$$

Vrednost L je definirana na intervalu $[0,100]$ za relativno svetlost (Y/Y_n) , ki je definirana na intervalu $[0,1]$. Obstajajo še trije izpeljani parametri v polarnem koordinatnem sistemu, ki določajo barvo, ki je bližje človeški predstavitvi: kromatičnost c , barvo tona h_{uv} in nasičenost s_{uv} . Ti trije parametri določajo barvna prostora kot sta CIE-*Lhs* in CIE-*Lhc*.

CIE-Lab

CIE-*Lab* barvni prostor ima vse lastnosti in značilnosti barvnega prostora CIE-*Luv*, s to razliko, da izhaja iz barvnega sistema CIE- XYZ in ima zaradi tega tudi drugačno transformacijo za vrednosti a in b . Vrednost L pa ostaja enaka.

$$\begin{aligned} a &= 500 * (f(X/X_n) - f(Y/Y_n)), \\ b &= 200 * (f(Y/Y_n) - f(Z/Z_n)), \end{aligned}$$

$$\text{kjer je } f(t) = \begin{cases} t^{\frac{1}{3}}; & \text{ce } t > 0.008856, \\ 7.787 * t + 16/116; & \text{ce } t \leq 0.008856. \end{cases}$$

Tudi pri tem barvnem prostoru imamo možnost boljše predstavitve barv s polarnimi koordinatami. V tem sistemu ne moremo izraziti vrednosti nasičenosti barve (ang. saturation), zato lahko izpeljemo samo barvni prostor CIE- Lch , kjer c določa kromatičnost in h barvo tona.

Ko določamo transformacijo na sistemih s katodno cevjo, CRT, običajno vzamemo za referenčno belo točko kar belo točko, ki je definirana na takšnem CRT sistemu. Barvna prostora CIE- Luv in CIE- Lab najboljše posnemata človeško dožemanje svetlobe in barv. K temu seveda pripomore zapletena pretvorba iz barvnega prostora RGB, kar vzame zelo veliko računskega časa. Kljub temu se ta dva prostora veliko uporabljata pri iskanju kožne barve na sliki. Barvna prostora CIE- Luv in CIE- Lab sta bila uporabljena v [79, 80, 84, 60]. Osnovni barvni prostor CIE- Yxy pa je v svojem delu uporabil Terrillon [69].

3.1.7 Razmerja med RGB kanali

Znano je, da barva kože vsebuje zelo veliko rdeče barvne komponente. Zaradi tega se za iskanje kožne barve, veliko uporablja razmerje med rdečim barvnim kanalom in ostalima barvnima kanaloma G in B . Razmerje R/G sta uporabila Wark in Sridharan [73]. Ostala razmerja, kot so R/G in G/B pa so bila preizkušena v delu Branda in Masona [5].

3.2 Modeliranje kožne barve

Glavni cilj iskanja obrazov na osnovi barv je zgraditi takšno odločitveno pravilo, ki bo razlikovalo med slikovnimi elementi barve kože in ostalih barv, ki predstavljajo "nekožno barvo". To je običajno povezano z merilom, ki meri oddaljenost slikovnega elementa izbrane barve od barve kože. V nadaljevanju bomo predstavili različne metode, s katerimi lahko določimo takšno merilo. Pri prvi metodi omejimo barvni prostor na podlagi pravil, pri drugih dveh pa zgradimo merilo na podlagi barvnega modela.

3.2.1 Eksplicitno določene meje

Najpreprostejša metoda za izgradnjo klasifikatorja je eksplicitno določanje mej, ki omejujejo področje kožne barve. Te meje določimo s pomočjo pravil. Področje znotraj teh

meja določi tako imenovani klaster ali kožni grozd. Tako definirana pravila smo uporabili tudi pri postopku izločanja kožne barve v interaktivni instalaciji: "15 sekund slave" [63].

Enostavnost te metode je pritegnila tudi veliko drugih raziskovalcev, ki so jo uporabili pri segmentaciji kožne barve [1, 19, 30]. Prednost te metode je prav njena enostavnost, ki omogoča izgradnjo zelo hitrega klasifikatorja. Glavna slabost pa je v tem, da moramo za dobro klasifikacijo izbrati ustrezen barvni prostor in natančno določiti pravila.

Da bi premagali ti dve slabosti, je bila predlagana metoda, ki s pomočjo strojnega učenja določi primeren barvni prostor in izbere optimalno pravilo [22]. Avtorji začnejo z normaliziranim barvnim prostorom RGB in nato postopno izvajajo indukcijski algoritem izgradnje (ang. Constructive induction algorithm), ki zgradi novo množico atributov. Ta množica se zgradi na podlagi normalizirane trojice r , g in b in konstantne vrednosti $1/3$ ter preprostih aritmetičnih pravil. Za vsako množico atributov nato določijo odločitveno pravilo, ki najbolje ločuje med obrazi in ne-obrazi. Avtorji preprečijo iskanje preveč kompleksnih pravil, ker bi to lahko vodilo k prevelikemu ujemanju z učno množico, ki se lahko pojavi tudi, če imamo opraviti s premalo reprezentativno učno množico. S pomočjo tako najdenih pravil so nad svojo testno množico dosegli boljše rezultate, kot z Baysovim klasifikatorjem.

3.2.2 Neparametrično modeliranje

S pomočjo neparametrične metode zgradimo porazdelitev kožne barve iz učnih primerov, brez predhodno izpeljanega eksplisitnega modela. Takšno porazdelitev največkrat imenujemo tudi SPM (ang. Skin Probability Map), ki vsaki diskretni vrednosti barvnega prostora pripiše verjetnost, da ta pripada kožni barvi.

Normalizirana tabela stanj (ang. Lookup table, LUT)

Veliko algoritmov za iskanje in sledenje človeškemu obrazu uporablja za segmentacijo obrazne barve metodo, ki temelji na barvnem histogramu [84, 29]. Barvni prostor oziroma njegov kromatični del se najprej razbije na enake dele. Vsak tako razdeljen del pripada določenemu barvnemu področju parov v 2D barvnem prostoru ali trojic v 3D barvnem prostoru. Tako zgrajen histogram v 2D ali 3D prostoru nato predstavimo s pomočjo tabele

stanj LUT (ang. lookup table). Posamezen del tabele vsebuje število, kolikokrat se je določena barva, ki pripada temu območju, pojavila v učni množici obrazov. Po končanem postopku gradnje histograma oziroma učenja, dobljeni histogram normaliziramo tako, da predstavlja vrednosti histograma diskretno verjetnostno porazdelitev:

$$P_{skin} = \frac{skin[c]}{Norm}, \quad (3.2)$$

kjer predstavlja $skin[c]$ število pojavljanja določene barve c v histogramu. $Norm$ je normalizacijska konstanta, določena kot vsota vseh vrednosti v histogramu [28] ali maksimalna vrednost v histogramu, ki pripada določenemu območju [84]. Tako normalizirana vrednost P_{skin} nam predstavlja verjetnost, da izbrana barva c pripada kožni barvi.

Bayesov klasifikator

Vrednost $P_{skin}(c)$ določena v enačbi 3.2 je pravzaprav pogojna verjetnost $P(c|skin)$. Določa vrednost verjetnosti opazovane barve c , ob prepostavki da je to kožna barva. Bolj primerno merilo za razločevanje kožne barve bi bila vrednost $P(skin|c)$, ki določa vrednost verjetnosti kožne barve pri določeni barvi c . Za izračun te verjetnosti uporabimo Bayesovo pravilo:

$$P(skin|c) = \frac{P(c|skin)P(skin)}{P(c|skin)P(skin) + P(c|\neg skin)P(\neg skin)}, \quad (3.3)$$

Vrednosti za verjetnosti $P(c|skin)$ in $P(c|\neg skin)$ izračunamo neposredno iz barvnega histograma (enačba 3.2). Apriorni verjetnosti $P(skin)$ in $P(\neg skin)$, pa lahko določimo iz celotnega števila vseh primerov obrazov in neobrazov učne množice [28, 84]. Kot pravilo za iskanje obrazne barve lahko uporabimo neenačbo $P(skin|c) \geq \Theta$, kjer je Θ mejna vrednost. Na podlagi pravila lahko nato izrišemo krivuljo ROC (ang. receiver operating characteristic). Krivulja ROC prikazuje razmerje med pravilno in napačno klasificiranimi barvnimi elementi kože za dano pravilo, kot funkcijo mejne vrednosti Θ . Iz krivulje lahko ugotovimo, da je klasifikacijsko pravilo $P(skin|c) \geq \Theta$ neodvisno od izbire apriornih verjetnosti. Ta ugotovitev izhaja že iz same narave Bayesovega klasifikatorja. Iz tega sledi, da izbira apriorne verjetnosti $P(skin)$ vpliva samo na izbiro mejne vrednosti Θ .

Izračunu vrednosti celotne verjetnosti po enačbi 3.3 se lahko izognemo, saj potrebujemo samo primerjavo med $P(\text{skin}|c)$ in $P(\neg\text{skin}|c)$, ki jo lahko zapišemo kot:

$$\frac{P(\text{skin}|c)}{P(\neg\text{skin}|c)} = \frac{P(c|\text{skin})P(\text{skin})}{P(c|\neg\text{skin})P(\neg\text{skin})}. \quad (3.4)$$

Primerjava enačbe 3.4 z mejno vrednostjo Θ nas pripelje do odločitvenega pravila za iskanje barve obraza oziroma neobraza. Po določenem preurejanju lahko odločitveno pravilo zapišemo kot:

$$\frac{P(\text{skin}|c)}{P(\neg\text{skin}|c)} > \Theta, \quad (3.5)$$

$$\Theta = K \times \frac{1 - P(\text{skin})}{P(\text{skin})} \quad (3.6)$$

To pravilo nam tudi pokaže, zakaj izbira apriornih verjetnosti ne vpliva na splošno obnašanje klasifikacijske metode, kajti za katero koli apriorno verjetnost $P(\text{skin})$ je mogoče izbrati primerno vrednost konstante K , ki nam bo dala enako mejno vrednost Θ .

Metoda SOM (ang. Self Organizing Map)

Samoorganizacijska mreža (SOM), ki jo je v osemdesetih odkril Kohonen je danes najpopularnejša mreža med nenadzorovanimi nevronskimi mrežami. V delu [7] so jo uporabili za izgradnjo klasifikatorja, ki ločuje barvo kože. Zgradili so dve mreži: prva predstavlja samo barvo kože, druga predstavlja tako barvo kože kot tudi druge barve. Mrežo so naučili s pomočjo učne množice 500 ročno označenih slik obrazov. Učinkovitost metode so preizkusili na učni in testni množici avtorjev ter na Compaqovi bazi obrazov [28]. Na SOM mreži so bili uporabljeni različni barvni prostori, kot so normaliziran prostor RGB, barvni prostor HS, kartezijski barvni prostor HS in kromatični del barvnega prostora TSL. Rezultati so pokazali, da uporaba različnih barvnih prostorov ne vpliva vidno na učinkovitost same metode. Metoda s pomočjo SOM mreže je na avtorjevi preizkusni množici primerov pokazala boljše rezultate kot pri uporabi Gaussovega modela, po drugi strani pa je bila na Compaqovi bazi slabša od metode na osnovi histograma barvnega prostora RGB. Avtorji [28] so tudi poudarili, da metoda na podlagi SOM mreže očitneje

manj obremenjeni vire v primerjavi z metodami na osnovi histograma oziroma na podlagi modela. Mreže SOM so primerne tudi za implementacijo v realnem času, na primerni strojni opremi, ki podpira takšne mreže.

Povzetek neparametričnih metod

Glavne prednosti neparametričnih metod so:

- hitrost pri učenju in med samo uporabo,
- teoretična neodvisnost od porazdelitve kožne barve.

Predvsem zadnja lastnost za ostali dve metodi (na osnovi pravil in parametričnih metod) ne velja. Slabosti neparametričnih metod so:

- za delovanje potrebujejo veliko prostora,
- ne omogočajo interpolacije ali generalizacije učnih podatkov.

Da zmanjšamo prostorsko zahtevnost metode in še vedno dovolj ločimo učne primere, lahko ustrezno zmanjšamo barvni prostor. Če imamo na primer barvni prostor RGB in vsako barvo predstavimo z 8 biti, moramo imeti $2^{24} = 16$ milijonov elementov, da shranimo vse verjetnosti za barvo kože. Preizkus različnih velikosti barvnega prostora RGB je pokazal, da je izbira: $32 \times 32 \times 32$ najprimernejša [28].

3.2.3 Parametrične metode

Metode, ki temeljijo na neparametričnih modelih histogramov potrebujejo za svoje delovanje veliko prostora, njihova uspešnost pa je močno odvisna od učne množice. Če potrebujemo bolj splošen model kožne barve, ki omogoča posploševanje in interpolacijo učni primerov, moramo uporabiti modele kožne porazdelitve, ki so odvisni od parametrov oziroma tako imenovane parametrične modele (ang. parametric models).

Enojni Gaussov model

Porazdelitev kožne barve lahko modeliramo s pomočjo eliptične Gaussove verjetnostne funkcije PDF (ang. probability density function), ki je definirana kot:

$$P(c|skin) = \frac{1}{2\pi|\Sigma_s|^{1/2}} \cdot e^{-\frac{1}{2}(c-\mu_s)^T \Sigma_s^{-1} (c-\mu_s)}. \quad (3.7)$$

Vhodni vektor c predstavlja barvo. Parametra porazdelitve μ_s in Σ_s označujeta vektor povprečnih vrednosti in kovariančno matriko. Parametra μ_s in Σ_s , izračunamo iz učne množice slik obrazov:

$$\mu_s = \frac{1}{n} \sum_{j=1}^n c_j, \quad \Sigma_s = \frac{1}{n-1} \sum_{j=1}^n (c_j - \mu_s)(c_j - \mu_s)^T \quad (3.8)$$

Vrednost n predstavlja celotno število vseh učnih primerov c_j . V našem primeru so to slikovni elementi, ki predstavljajo kožno barvo. Izračunano verjetnost $P(c|skin)$ lahko uporabimo neposredno, kot merilo kako podobna je izbrana barva c kožni barvi [45]. Ekvivalentno pa lahko kot merilo uporabimo tudi razdaljo vektorja c od vektorja povprečij μ_s ob upoštevanju kovariančne matrike Σ_s [69]:

$$\lambda_s = (c - \mu_s)^T \Sigma_s^{-1} (c - \mu_s). \quad (3.9)$$

Ta razdalja se imenuje Mahalanobisova razdalja (ang. Mahalanobis distance). Enojni Gaussov model so v svojih delih uporabili [25, 79].

Mešanica Gaussovih modelov

Mešanica Gaussovih modelov (ang. Mixture of Gaussians) je sestavljena iz več Gaussovih modelov in omogoča modeliranje kompleksno porazdeljenih struktur. Je posplošitev navadnega enojnega Gaussovega modela. Verjetnostno funkcijo porazdelitve v tem primeru lahko zapišemo kot:

$$P(c|skin) = \sum_{i=1}^k \pi_i \cdot P_i(c|skin), \quad (3.10)$$

kjer predstavlja k število vseh mešanih komponent. Mešani parametri so predstavljeni z vrednostjo π_i , katerih vsota je normalizirana konstanta: $\sum_{i=1}^k \pi_i = 1$. Vrednosti $P_i(c|skin)$ pa predstavljajo posamezne funkcije Gaussove porazdelitve, vsaka s svojo povprečno vrednostjo in kovariančno matriko. Učenje samega modela in s tem določitev parametrov modela določimo z dobro znanim iterativnim postopkom, znanim kot EM (ang. Expectation-Maximization) algoritem. Ta algoritem predpostavlja vnaprej znano število mešanih komponent k . Sam postopek EM algoritma je opisan v [79, 69]. Klasifikacijo s pomočjo mešanice Gaussovih modelov nato določimo s primerjavo izračunane verjetnosti $P(c|skin)$ z določeno mejno vrednostjo.

Zelo pomembna je pravilna izbira števila komponent k . Po eni strani mora model dovolj dobro opisovati učno množico, po drugi pa ne sme biti preveč kompleksen, da se ne prilagodi preveč na učno množico (ang. over-fitting). Število uporabljenih komponent se od raziskovalca do raziskovalca zelo razlikuje. Nekateri uporabljajo samo 2 komponenti, nekateri pa 16 komponent. S pomočjo preizkušanja sta avtorja Yang in Ahuja [79] izbrala 2 komponenti. Kot dober kompromis med točnostjo in kompleksnostjo pa je Terrillon [69] izbral 8 komponent. Mešanico Gaussovih modelov so uporabili tudi drugi avtorji [43, 47].

Model Eliptičnih mej

S preučevanjem porazdelitev kožnih in nekožnih barv v različnih barvnih prostorih, sta avtorja Lee in Yoo [37] ugotovila, da eliptični kožni skupek z Gaussovim modelom ni dovolj dobro predstavljen. Zaradi nesimetrične oblike kožnega skupka glede na konico krivulje, imamo pri simetričnem Gaussovim modelu veliko napačno najdenih kožnih elementov. Zato sta predlagala drugačno metodo, ki uporablja tako imenovani model eliptičnih mej (ang. elliptical boundary model). Metoda je enako hitra in enostavna za učenje in uporabo kot metoda, ki temelji na Gaussovim modelu. Model eliptičnih mej se je izkazal za boljšega na Compaqovi bazi v primerjavi z enojnim Gaussovim modelom in mešanico Gaussovih modelov. Model eliptičnih mej je definiran kot:

$$\Theta(c) = (c - \phi)^T \Lambda^{-1} (c - \phi). \quad (3.11)$$

Postopek učenja je sestavljen iz dveh korakov. V prvem koraku zavržemo več kot 5% vseh učnih primerov, ki so redkeje zastopani v učni množici. S tem se znebimo

nepomembnih in šumnih podatkov. V drugem koraku določimo parametra ϕ in Λ , s pomočjo enačb:

$$\phi = \frac{1}{n} \sum_{i=1}^n c_i, \quad \Lambda = \frac{1}{N} \sum_{i=1}^n f_i \cdot (c_i - \mu)(c_i - \mu)^T, \quad (3.12)$$

$$\mu = \frac{1}{N} \sum_{i=1}^n f_i c_i, \quad N = \sum_{i=1}^n f_i. \quad (3.13)$$

Paziti moramo, ker vrednost n ni število vseh barv v učni množici ampak predstavlja število vseh značilnih barv c_i v celotni učni množici slikovnih elementov kožne barve. Vrednost f_i je število primerov posameznega barvnega vektorja c_i . Barvni element z določeno barvo c je tako klasificiran kot barvni element kože, če $\phi(c) < \Theta$, kjer je Θ neka mejna vrednost. Avtorja trdita, da je njun model boljši, ker nesimetričnost podatkov učne množice ne vpliva na izračun srednje vrednosti modela ϕ .

Povzetek parametričnih metod

Vedno, ko imamo opraviti z modeli, se pojavi vprašanje točnosti oziroma pristnosti modela. Očitno je, da je pri parametričnih metodah porazdelitev kožne barve in s tem tudi prava izbira barvnega prostora bolj pomembna, kot pri neparametričnih metodah. Pri uporabi različnih barvnih prostorov se uspešnost metode na podlagi parametričnega modela zelo spreminja [69, 37].

Nekateri avtorji so v svoja dela vključili tudi teoretično razlago, zakaj so uporabili določen model v svojih raziskavah. Yang [77] je pokazal, da porazdelitev kožne barve človeka pri nespremenjeni svetlobi, v normaliziranem barvnem prostoru RGB, ustreza Gaussovemu modelu. Avtorja Yang in Ahuja [79] sta s pomočjo statističnega testa potrdila hipoteze o pravilnosti elementov kože v CIE-*Luv* barvnem prostoru in ustreznosti mešanice dveh Gaussovih modelov za modeliranje kožne barve.

Ostali avtorji se zanašajo predvsem na obliko kromatičnega skupka kožne barve v izbranem barvnem prostoru. Če je kožni skupek približno eliptične oblike uporabijo enojni Gaussov model ali kaj podobnega, če pa gre za neeliptično obliko se odločijo, da bodo skupek poskušali opisati z večjim številom Gaussovih modelov oziroma s pomočjo modela,

ki uporablja mešanico Gaussovih modelov. Nato na podlagi uspešnosti modela določijo, ali je model sprejemljiv ali ne [37, 43, 69].

3.3 Izbira prave metode in ustreznega barvnega prostora

Za segmentacijo kožne barve lahko uporabimo različne metode na različnih barvnih prostorih. Najenostavnejša metoda eksplicitno določenih mej temelji na enostavnih pravilih. Barvo kože pa lahko predstavimo tudi s pomočjo modela, in sicer s pomočjo parametrov kot parametrični model, ali pa s pomočjo histograma kot neparametrični model. Na podlagi učne množice in želenih rezultatov se moramo odločiti za najprimernejšo metodo. Pri reprezentativno veliki učni množici, kjer nam čas in kompleksnost metode nista pomembna, izberemo neparametrično metodo. Pri majhni učni množici, kjer nas malce večja napaka nepravilno odkritih barvnih elementov ne moti in nam je pomembnejša hitrost, pa se odločimo za parametrično metodo.

Ustrezen barvni prostor izberemo na podlagi metode, ki smo jo izbrali. Barvni prostor je pomembnejši pri parametričnih modelih, kjer je ugodno, da tvori barvni skup v izbranem barvnem prostoru. Za metodo na osnovi pravil in metodo na osnovi neparametričnega modela pa barvni prostor ni ključnega pomena.

3.3.1 Povzetek različnih metod

Glavna prednost metod, ki uporabljajo eksplicitno določene meje (razdelek 3.1), je njihova preprostost in intuitivnost klasifikacijskih pravil. Njihova težava pa je tako v izbiri pravega barvnega prostora, kot tudi v izbiri primernih odločitvenih pravil. Predlagana je bila metoda, ki s pomočjo strojnega učenja poišče ustrezen barvni prostor in preprosto odločitveno pravilo [21]. S pomočjo te metode lahko omilimo omenjeno težavo teh metod.

Neparametrične metode (razdelek 3.2) so hitre tako pri učenju kot pri klasificiranju, neodvisne so od oblike porazdelitve in posledično tudi od barvnega prostora. Po drugi strani pa potrebujejo za svoje delovanje veliko podatkovnega prostora in veliko reprezentativno učno množico.

Parametrične metode (razdelek 3.2.3) so hitre, imajo pa še to dobro lastnost, da omogočajo interpolacijo in s tem posplošitev nepopolne učne množice. Predstavlja jih nekaj parametrov in za delovanje potrebujejo zelo malo podatkovnega prostora. Če uporabljamo mešanico Gaussovih modelov, so lahko parametrične metode tudi počasne in to tako pri učenju, kot pri uporabi. Močno so odvisne od oblike porazdelitve učnih primerov. Poleg tega veliko parametričnih metod uporablja za klasifikacijo samo model kožne barve, ne pa modela nekožne barve. Posledično imamo zaradi tega, v primerjavi z neparametričnimi metodami, veliko napačno odkritih slikovnih elementov FP (False Positives).

3.3.2 Povzetek različnih barvnih prostorov

Na prvi pogled se nam zdi izbira pravega barvnega prostora ključnega pomena za iskanje kože s pomočjo barve, vendar temu ni tako. Poleg barvnega prostora sta pomembna tudi tehnika modeliranja ter obsežnost problema, na katerem iščemo izbrano barvo. Pojavlja se vprašanje: kateri je najboljši barvni prostor za segmentacijo kožne barve? Ali bolj splošno: ali sploh obstaja optimalni barvni prostor za barvno klasifikacijo obraza? Presenetljivo veliko del, ki se ukvarja s segmentacijo na osnovi barve, ne omenja, zakaj je bil izbran določen barvni prostor. Verjetno zato, ker bi izbira katerega koli barvnega prostora, na njihovem omejenem problemu oziroma omejeni učni množici, dal enake ali podobne rezultate.

Kar nekaj del se ukvarja s problematiko izbire pravega oziroma najboljšega barvnega prostora [84, 69, 21, 22]. Nekateri avtorji, ki se v svojih delih zavedajo pomembnosti izbire pravega barvnega prostora, zato preizkusijo več barvnih prostorov [79, 77, 60].

Ustreznost barvnega prostora se ocenjuje na različne načine. Prva metoda, s pomočjo katere lahko ocenimo napako, izračuna napako med učno in testno množico na podlagi zgrajenega modela. Ta ocena nam jasno pove, kako dobro se model prilega dani množici. Druga metoda temelji na prekrivanju barvnih elementov, ki pripadajo obrazu in elementov, ki ne pripadajo obrazu. Naslednja metoda ocenjuje, kako kompaktno so porazdeljeni barvni elementi kože v izbranem barvnem prostoru. Vse te ocene nam dajo nek splošni pregled na porazdelitev barvnih elementov kože in nekože. Pravega merila, ki bo določilo najboljši barvni prostor, na takšen način ne moremo dobiti. Izbira pravega barvnega pros-

tora je v veliki meri odvisna od modela, ki smo ga predhodno izbrali. In na podlagi tega modela lahko nato izberemo tudi najustreznejši barvni prostor. Izbira pravega prostora je pri modeliranju izrednega pomena, kar tudi kažejo raziskave [69, 37]. Tudi metode, ki temeljijo na pravilih [63, 21], so odvisne od barvnega prostora. Neparametrične metode, kot so Bayes, SPM, SOM, LUT, pa so v glavnem neodvisne od barvnega prostora. Na te metode izbira barvnega prostora v glavnem nima vpliva [84, 7]. Albiol je to tudi teoretično dokazal [2]. Seveda pa prekrivanje med slikovnimi elementi barve kože in nekože močno vpliva na točnost parametričnih metod in metod na podlagi tabel, ker njihovi modeli tega prekrivanja ne upoštevajo.

Veliko raziskovalcev, s področja iskanja obrazov na podlagi barv, iz barvnega prostora odstrani komponento svetlosti (ang. *luminance*). Takšna izbira se zdi logična, kajti cilj je modelirati barvo kože, ki je bolj določena z barvnimi komponentami kot s svetlostjo barve. Neupoštevanje svetlosti barve tudi posledično zmanjša dimenzijo barvnega prostora, kar je v veliki meri povezano z enostavnejšo in hitrejšo metodo iskanja obraza. Razlog za neupoštevanje svetlosti je tudi v tem, da se človeška koža od človeka do človeka razlikuje predvsem po svetlosti in manj po barvi sami. Različne osvetlitve zelo vplivajo na barve, zato je cilj vseh metod, ki temeljijo na barvah, da so čimbolj neodvisne od osvetlitve. S pomočjo odstranitve svetlosti barve to delno dosežemo. Upoštevanje samo kromatičnosti barve je še posebej dobrodošlo, kadar imamo učno množico, v kateri se pojavljajo zelo majhne razlike v osvetlitvi. V tem primeru lahko zgradimo model, ki bo pravilno klasificiral tudi barve kože z drugačno osvetlitvijo od učne množice.

3.3.3 Izbira primerne barvnega prostora

Na podlagi raziskav različnih barvnih prostorov za segmentiranje kožne barve [69] so se najbolj izkazali kromatični barvni prostori (rg , HS , TS , $CrCb$, xy , itd). To so tisti barvni prostori, ki upoštevajo samo kromatični del barvnega prostora in izpustijo svetlost barve. Izkazalo se je tudi, da so za modeliranje s parametričnim modelom najprimernejši normalizirani barvni prostori (rg , TS , xy , itd). Neparametrični model je neodvisen od izbire barvnega prostora, zato so predlagani normalizirani kromatični barvni prostori ravno tako ustrezni tudi za modeliranje s pomočjo neparametričnih modelov. Med izbranimi barvnimi prostori smo izbrali barvni prostor rg , ki je najenostavnejši in najhitrejši

pri pretvorbi iz barvnega prostora RGB.

3.3.4 Preizkus različnih metod

Kožno barvo smo poskušali modelirati s pomočjo dveh različnih modelov. S pomočjo parametričnega modela in s pomočjo neparametričnega modela. Za parametrični model smo izbrali Gaussov model, ki je najbolj ustrežal porazdelitvi učne množice primerov. Za neparametrični model smo uporabili 2D histogram porazdelitve kožne barve. Neparametrični model smo nadgradili z dodatnim modelom nekožne barve, s pomočjo katerega smo poskušali izločiti nepravilno izbrane barve obraza.

V nadaljevanju bom podrobneje predstavil metodi na osnovi parametričnega in neparametričnega modela. Najprej pa bom predstavil metodo eksplicitno določenih mej, s katero bom primerjal rezultate obeh metod na osnovi modelov.

Poglavje 4

Implementacija različnih metod v izbranem barvnem prostoru

V nadaljevanju bomo predstavili tri različne metode za modeliranje kožne barve. Najprej bo predstavljena najenostavnejša metoda, ki jo trenutno uporabljamo v instalaciji "15 sekund slave" [63]. Metoda deluje pod določenimi svetlobnimi pogoji zelo dobro. Glavne slabosti te metode so, da je zelo občutljiva na osvetlitev, da prepozna samo kožo svetle polti in da včasih nepravilno izbere tudi barvo, ki ni koža. S pomočjo drugih dveh metod, ki temeljita na kromatičnem barvnem prostoru bomo poskušali omenjene probleme odpraviti ali vsaj omiliti. Poskusili bomo modelirati kožno barvo pod različnimi osvetlitvami in različnih kožnih polti. Medtem, ko pri drugi metodi modeliramo kožno barvo s pomočjo parametričnega modela, uporabljamo pri tretji metodi neparametrični model. Tretjo metodo lahko še dodatno izboljšamo s pomočjo modela nezaželenih barv in s tem še dodatno zmanjšamo nepravilne izbire. Na koncu bomo na podlagi preizkusov na različnih bazah slik izbrali najustreznejšo.

4.1 Metoda eksplicitno določenih mej

Metoda eksplicitno določenih mej temelji na preprostih pravilih v RGB barvnem prostoru. Metoda vsebuje dve skupini pravil: prva skupina določa omejitve pri normalni osvetlitvi, druga pa pri uporabi bliskavice ali pri majhni osvetlitvi.

Trojica (R,G,B) je klasificirana kot barva kože, če:

1. **pravilo:** $R > 95$ in $G > 40$ in $B > 20$ in
2. **pravilo:** $\max\{R, G, B\} - \min\{R, G, B\} > 15$ in
3. **pravilo:** $|R - G| > 15$ in
4. **pravilo:** $R > G$ in $R > B$.

Prvo pravilo omeji posamezne barvne kanale in določa barvo kože pri enakomerni osvetlitvi. Drugo pravilo določa, da posamezne komponente RGB ne smejo biti blizu skupaj - na ta način izločimo sivine. Tretje pravilo določa, da tudi barvna kanala R in G ne smeta biti blizu skupaj - s tem se omejimo samo na svetlo polt. Zadnje, četrto pravilo pa določi, da je R največja komponenta, oziroma, da v barvi kože prevladuje rdeča komponenta.

Za klasifikacijo kože ob uporabi bliskavice ali pri majhni osvetlitvi pa veljajo sledeča pravila:

- 1.**pravilo:** $R > 220$ in $G > 210$ in $B > 170$ in
- 2.**pravilo:** $|R - G| \leq 15$ in
- 3.**pravilo:** $R > G$ in $G > B$.

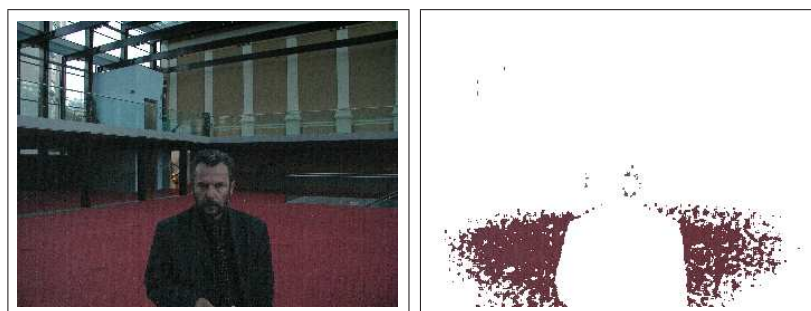
Prvo pravilo navzdol omeji barvni prostor RGB, drugo pravilo določi, da sta komponenti R in G blizu skupaj, tretje pravilo pa, da je B najmanjša komponenta.

Pravila uspešno segmentirajo kožno barvo svetle polti pri normalni osvetlitvi. Težave pa se pojavijo pri nestandardnih osvetlitvah. V takem primeru metoda izbere za kožno barvo tudi druge barve (slika 4.1) ali pa kožne barve ne prepozna pravilno (slika 4.2).

Predstavljena pravila imajo še to omejitev, da pravilno najdejo samo kožno barvo svetle polti. Omenjene probleme je zelo težko odpraviti s spreminjanjem obstoječih pravil (t.j. s pomočjo dodatnih pravil ali posplošitvijo trenutnih pravil), zato smo se odločili, da poskusimo izločiti kožno barvo iz slike s pomočjo metod na osnovi barvnega modela. Za modeliranje smo najprej uporabili parametrični in nato še neparametrični model.



Slika 4.1: Nepravilno ohranjena barva.



Slika 4.2: Kožna barva, ki zaradi nestandardne osvetlitve ni bila prepoznana.

4.2 Metoda na osnovi parametričnega modela

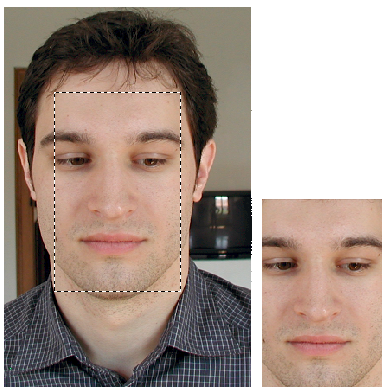
Za modeliranje kožne barve smo uporabili kromatični barvni prostor rg , ki se je med preizkušenimi barvnimi prostori: CIE_{xy}, CIE-dsh, HSV, YIQ, YES, CIE_{Luv}, CIE-Lab in TSL izkazal za zelo uspešnega [69]. Kljub temu, da se je za najboljši barvni prostor izkazal TSL barvni prostor, smo se odločili za preprostejši rg barvni prostor, ki po uspešnosti ne zaostaja veliko, pa še pretvorba iz RGB barvnega prostora je enostavnejša:

$$r = \frac{R}{R + G + B}, \quad g = \frac{G}{R + G + B}. \quad (4.1)$$

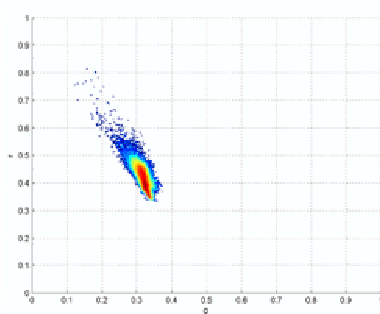
S pomočjo te pretvorbe tudi zmanjšamo dimenzijo barvnega prostora iz tridimenzionalnega v dvodimenzionalni prostor ($3D \rightarrow 2D$). Porazdeljenost kožne barve lahko zaradi tega opišemo s pomočjo dvodimenzionalnega barvnega histograma.

Barvni histogram opisuje porazdeljenost barv v barvnem prostoru in se veliko uporablja

pri analizi barvnih slik v računalniškem vidu. Že sredi osemdesetih let 20 stoletja so odkrili, da tvori barvni histogram osvetljene nehomogene površine v 3D barvnem prostoru ravnino [78]. Dokazano je bilo tudi, da se barve ne porazdelijo naključno po ravnini ampak tvorijo barvne skupke v določenih točkah. Te ugotovitve potrjuje tudi histogram kožne barve. Slika 4.3 prikazuje obraz in njegov izrez kožne barve. Porazdelitev izrezane kožne regije predstavimo s pomočjo 2D histograma v barvnem prostoru rg (slika 4.4). Vidimo lahko, da kožna barva predstavlja samo majhen del celotnega kromatičnega barvnega prostora rg oziroma povedano drugače, samo majhen del vseh barv predstavlja kožno barvo.



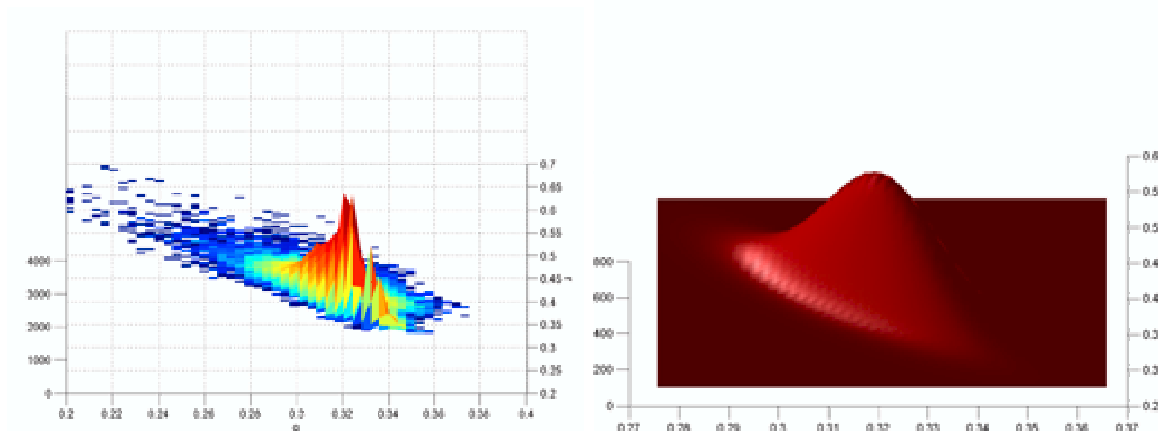
Slika 4.3: Primer obraza in označene kožne regije.



Slika 4.4: Porazdelitev kožne barve v kromatičnem barvnem prostoru.

Če si porazdelitev kožne barve ogledamo bližje vidimo, da tvori histogram prav posebno

obliko. Opazimo lahko, da oblika porazdelitve sovpada z Gaussovo porazdelitvijo (slika 4.5).



Slika 4.5: Porazdeljenost kožne barve in pripadajoč Gaussov model.

Torej lahko porazdelitev opišemo s pomočjo enojnega Gaussovega modela $N(\mu, \Sigma)$ dveh spremenljivk:

$$N(\mu, \Sigma) = \frac{1}{2\pi\sqrt{|\Sigma|}} e^{-\frac{1}{2}\lambda^2}, \quad (4.2)$$

kjer λ označuje Mahalanobisovo razdaljo med slikovnim elementom $c(i, j)$ in povprečno vrednostjo $\mu(i, j) = [r(i, j) \ g(i, j)]^t$:

$$\lambda^2 = [c - \mu]^t \Sigma^{-1} [c - \mu]. \quad (4.3)$$

Enačba 4.3 določa eliptično površino v kromatičnem prostoru velikosti λ in centrom v točki μ , eliptične osi pa določa kovariančna matrika Σ .

Povprečno vrednost $\mu = (r, g)$ in kovariančno matriko Σ pa izračunamo:

$$r = \frac{1}{n} \sum_{i=1}^n r_i, \quad g = \frac{1}{n} \sum_{i=1}^n g_i, \quad (4.4)$$

$$\Sigma = \frac{1}{n-1} \sum_{j=1}^n (c_j - \mu)(c_j - \mu)^t, \quad (4.5)$$

kjer vrednost n označuje število vseh slikovnih elementov.

Postopek za izgradnjo kožnega modela je sledeč:

1. izberemo obraz ali pa množico obrazov, če potrebujemo splošni model,
2. ročno izločimo kožno regijo (slika 4.3),
3. izračunamo povprečno vrednost μ in kovarinačno matriko Σ kožne porazdelitve v kromatičnem barvnem prostoru (enačbi 4.4 in 4.5),
4. izračunani vrednosti uporabimo za določitev Gaussove porazdelitve in s tem Gaussovega modela (slika 4.5),
5. ker ima model samo šest različnih parametrov, ga je enostavno prilagoditi različnim pogojem osvetlitve oziroma različnim kožnim barvam ljudi.

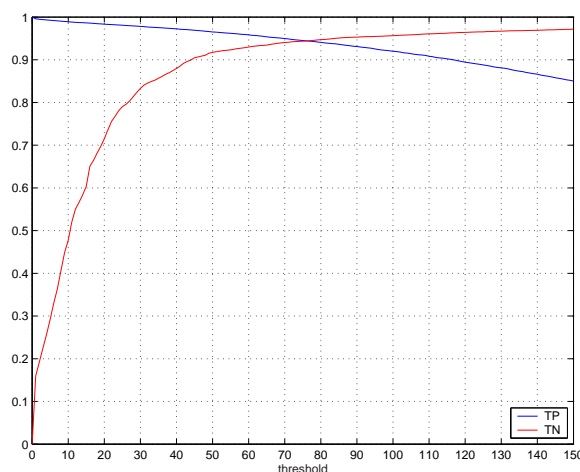
Da izločimo kožno barvo iz vhodne slike, jo moramo primerjati z izračunanim modelom kožne barve. Najprej vsak slikovni element vhodne slike pretvorimo v kromatični barvni prostor in ga primerjamo s porazdelitvijo kožne barve v modelu. Ujemanje dobimo s pomočjo izračunane verjetnosti Gaussove funkcije:

$$p(c|skin) = \frac{1}{2\pi|\Sigma|^{1/2}} e^{-\frac{1}{2}(c-\mu)^t \Sigma^{-1}(c-\mu)} \quad (4.6)$$

Večja kot je vrednost p za izbrano barvo c , večja je verjetnost, da barva c pripada kožni barvi. Tako izračunano verjetnost p nato primerjamo z mejno vrednostjo Φ_p . Če je vrednost večja od mejne vrednosti, izbrana barva c pripada kožni barvi, drugače pa ne.

Mejno vrednost Φ_p določimo s pomočjo učne množice slik obrazov in učne množice slik, na katerih ni kožne barve (t.i. množice neobrazov). Za učno množico obrazov lahko uporabimo kar učno množico na podlagi katere smo zgradili model, to je približno 10 slik velikosti 200×200 slikovnih elementov. Za učno množico neobrazov pa vzamemo različna ozadja obrazov, ki seveda ne smejo vsebovati kožne barve. Velikost učne množice naj bo vsaj toliko velika kot učna množica obrazov. Na izbrani množici slik obrazov dobimo razmerje med pravilno izbranimi slikovnimi elementi TP (ang. True Positives) in vsemi slikovnimi elementi. Na množici slik ozadij, ki ne vsebujejo kožne barve pa dobimo

razmerje med pravilno izločenimi slikovnimi elementi TN (ang. True Negatives) in vsemi slikovnimi elementi. Mejno vrednost Φ_p določa tista vrednost, ko sta razmerji TP in TN enaki: $TP = TN$ ali ekvivalentno, ko sta enaki razmerji FN za slike obrazov in FP za slike neobrazov: $FN = FP$. Slika 4.6 prikazuje funkciji za pravilno označene slikovne elemente (TP) in pravilno izločene slikovne elemente (TN) pri različnih mejnih vrednostih v intervalu $[0, 150]$. Funkciji se sekata pri mejni vrednosti 76, ko sta razmerji TP in TN enaki vrednosti 94%.



Slika 4.6: Grafa funkcij TP in TN za določitev mejne vrednosti Φ_p

Tako določeno mejno vrednost Φ_p lahko sedaj uporabimo za razločevanje kožne barve na poljubni vhodni sliki. Slikovne elemente vhodne slike primerjamo z mejno vrednostjo in če je vrednost Gaussove funkcije večja od mejne vrednosti: $p(c(i, j)|skin) > \Phi_p$, označimo slikovni element na položaju (i, j) z vrednostjo 1 (bela barva), drugače pa z 0 (črna barva). Na takšen način dobimo binarno sliko, na kateri označuje bela barva kožno barvo in črna vse ostale barve (slika 4.7).

Seveda pa je uspešnost te metode zelo odvisna od učne množice na podlagi katere je bil zgrajen barvni model. Če vhodna slika pripada isti podatkovni bazi, kot učna množica ali pa ustreza podobnim pogojem osvetlitve, so rezultati boljši. V nadaljevanju bomo metodo preizkusili na šestih različnih bazah barvnih slik, katerih slike se razlikujejo po osvetlitvi.



Slika 4.7: Določitev kožne barve na osnovi barvnega modela: a) vhodna slika, b) binarna slika z ohranjenimi področji kožne barve.

Na podlagi učnih primerov posamezne baze bomo za vsako bazo zgradili svoj model in s tako zgrajenim modelom poskušali razločevati kožno barvo na slikah iz baze na kateri je bil model naučen. Rezultate bomo primerjali z rezultati trenutne metode na osnovi eksplicitno določenih mej. Na koncu bomo poskusili zgraditi splošen model, ki bo upošteval učne primere vseh podatkovnih baz in njegovo uspešnost preizkusili na slikah iz vseh podatkovnih baz.

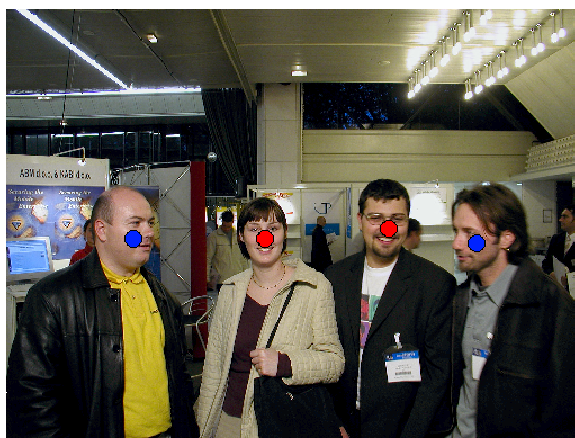
4.2.1 Postopek preizkusa uspešnosti metode

Če bi hoteli natančno določiti uspešnost metode na osnovi barve, bi morali ročno označiti slikovne elemente kožne barve vseh slik. Ker je takšen postopek zelo dolgotrajen in zamuden smo se odločili, da določimo uspešnost metod na drugačen način.

Na vseh testnih slikah smo ročno označili obraze in sicer ločeno smo označili obraze posnete od spredaj (rdeči krožec) in obraze posnete s strani (modri krožec) (slika 4.8) [72].

Nato smo s pomočjo metode izločili vse barve, ki niso v barvi kože. Ohranjeno kožno barvo smo združili v regije s pomočjo metode "rast regij" (slika 4.9b). Tako označene regije so predstavljale kandidate za obraz. Na koncu smo s preprostimi hevrističnimi pravili preverili kandidate, če ustrezajo karakteristikam obraza. Kandidat je bil izbran za obraz (slika 4.9c), če je ustrezal sledečim pravilom:

1. **pravilo:** delež kožne barve znotraj regije $> 40\%$,
2. **pravilo:** višina/širina regije ≤ 2.1 ,

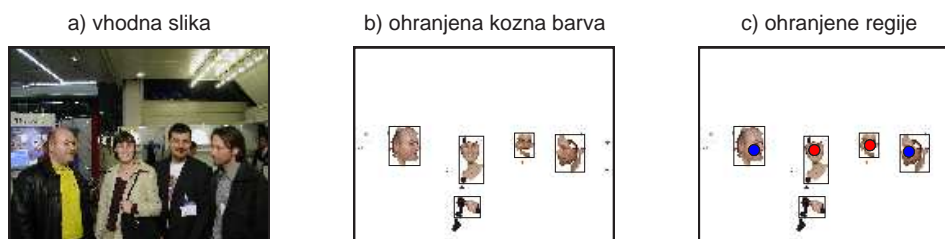


Slika 4.8: Primer slike z ročno označeni obrazi.

3. **pravilo:** širina/višina regije ≤ 1.6 ,
4. **pravilo:** minimalna širina regije glede na velikost slike = 5%,
5. **pravilo:** minimalna višina regije glede na velikost slike = 9%,
6. **pravilo:** maksimalna širina regije glede na velikost slike = 66%,
7. **pravilo:** maksimalna višina regije glede na velikost slike = 45%.

S pomočjo tako definiranih pravil izločimo veliko napačno ohranjenih kožnih regij. S pomočjo prvega pravila, ki določa, da mora imeti ohranjena regija več kot 40% kožne barve izločimo predvsem roke, ki so v določenem položaju v ustreznem razmerju, ki ju določata drugo in tretje pravilo. S pomočjo četrtega in petega pravila omejimo velikost regij navzdol in s pomočjo zadnjih dveh pravil omejimo velikost regij navzgor. Do posameznih vrednosti, ki jih uporabljamo v pravilih smo prišli s preizkušanjem mejnih vrednosti, ki še zadoščajo karakteristikam obraza.

Za vse uspešno najdene obraze smo nato preverili, če se središče predhodno označenega obraza nahaja znotraj ohranjene regije. V primeru, če je ohranjena regija vsebovala oznako obraza, smo povečali število pravilno najdenih obrazov TP, če pa regija ni vsebovala oznake, smo povečali število napačno najdenih obrazov FP (slika 4.9c).



Slika 4.9: Postopek preverjanja uspešnosti metode.

Kljub temu, da nam takšen način ni dal točne mere za uspešnost metode, pa nam je zelo dobro služil za primerjavo med različnimi metodami na osnovi barve.

V nadaljevanju bomo s pomočjo opisanega postopka primerjali med seboj različne metode na osnovi barve, kot so metoda eksplicitno določenih mej, parametrična in neparametrična metoda.

4.2.2 Modeliranje kožne barve pod različnimi osvetlitvami

Za preizkus modeliranja kožne barve pod različnimi osvetlitvami smo uporabili šest različnih podatkovnih baz z barvnimi slikami različnih osvetlitev. Slike so bile zajete pod različnimi osvetlitvami v različnih prostorih in z različnimi nastavitvami na sami napravi za zajemanje slik (sliki 4.10 in 4.11).

Slike posameznih podatkovnih baz smo dobili s pomočjo instalacije 15 sekund slave [63] in sicer na treh različnih lokacijah. Slike iz prvih treh baz so bile zajete na 8. mednarodnem festivalu računalniške umetnosti v Mariboru. Slike iz naslednjih dveh baz (4. baza in 5. baza) so bile zajete na prireditvi InterINFOS v Cankarjevem domu v Ljubljani. Zadnja baza slik pa je nastala v Finžgarjevi galeriji v Ljubljani.

Pri vseh bazah so viri osvetlitve zelo različni. Prva baza vsebuje slike, ki so bile zajete pod naravno osvetlitvijo. Pri drugi bazi je bil vir osvetlitve reflektor, pri tretji bazi pa je bil dodan svetlobni izvor vir iz lestenca, ki je oddajal rumeno svetlobo. Značilnosti prostora, v katerem so bile zajete slike prvih treh baz je, da so bile stene bele barve. Težave pri določanju kožne barve so se pojavile pri umetni osvetlitvi z rumeno svetlobo, kjer so postale stene zelo podobne barvi kože. Moteča so bila tudi lesena vrata v ozadju, ki so bila pod določenimi svetlobnimi pogoji barvno zelo podobna kožni barvi. Naslednji



Slika 4.10: Primer slike iz različnih podatkovnih baz.



Slika 4.11: Primer istega obraza pod različnimi osvetlitvami.

dve podatkovni bazi: četrta baza in peta baza sta bili zajeti na razstavnem prostoru sejma InterINFOS v Cankarjevem domu v Ljubljani. Pri četrti bazi je bil vir osvetlitve naravna svetloba iz ozadja in dodatne luči od zgoraj ter reflektor od spredaj. Nastavitev beline digitalnega fotoaparata je bila nastavljena na vir žarnice (ang. tungsten), zaradi tega so tudi slike zajete pod temi nastavitvami modrikaste. Pri slikah iz pete podatkovne baze pa je bil vir svetlobe izključno svetloba reflektorja. Značilnosti slik v peti podatkovni bazi so, da so slike zajete brez dodatne naravne svetlobe in so obrazi na slikah velikokrat zajeti iz profila. Zaradi slabih svetlobnih pogojev so slike v tej bazi pogosto neostre. Slike zadnje šeste podatkovne baze so bile zajete v galeriji, kjer je bila osvetlitev zelo slaba. Svetloba je prihajala iz treh različnih virov navadnih žarnic postavljenih nad obrazi. Zelo moteče so sence na obrazu, poleg tega so obrazi pogosto tudi premalo osvetljeni.

V tabeli 4.1 so predstavljene še ostale karakteristike posamezne podatkovne baze barvnih slik obrazov, kot so: število slik v podatkovni bazi, število vseh obrazov in število frontalnih obrazov na slikah.

	kraj	osvetlitev	št. slik	št. vseh obrazov	št. frontalnih obrazov
1. baza	Maribor	naravna	1027	2257	1526
2. baza	Maribor	reflektor	130	447	299
3. baza	Maribor	lestenec	63	140	112
4. baza	INFOS	naravna, luči, lestenec	3647	8476	3665
5. baza	INFOS	luči	451	970	381
6. baza	galerija	šibka svetloba luči	1261	3862	1166

Tabela 4.1: Lastnosti posameznih podatkovnih baz barvnih slik obrazov.

Na podlagi ročno izbranih obrazov smo za vsako posamezno podatkovno bazo izdelali svoj barvni model (slika 4.12). Učna množica posamezne podatkovne baze je vsebovala približno 10 različnih obrazov povprečne velikosti 200×200 slikovnih točk. Na sliki 4.12 lahko vidimo, da so si modeli med seboj zelo podobni in da vsi tvorijo majhen barvni skupek v kromatičnem barvnem prostoru.

Na osnovi tako zgrajenih modelov smo preizkusili uspešnost metode na posamezni podatkovni bazi slik iz katere smo zgradili model. Medtem ko smo model učili na podlagi 10 naključno izbranih reprezentativnih slik iz podatkovne baze, smo metodo preizkusili na vseh slikah izbrane podatkovne baze. Rezultate smo primerjali z metodo eksplicitno določenih mej (metoda 1) opisano v poglavju 4.1. Rezultate primerjave prikazuje tabela 4.2.

	1. baza		2. baza		3. baza		4. baza		5. baza		6. baza	
metoda	TP	FP	TP	FP	TP	FP	TP	FP	TP	FP	TP	FP
1.metoda	76%	35%	55%	31%	28%	33%	69%	13%	60%	12%	39%	9%
2.metoda	84%	26%	59%	21%	59%	36%	79%	15%	71%	17%	60%	15%

Tabela 4.2: Preizkus metode na osnovi različnih parametričnih modelov na slikah različnih osvetlitev.

Iz tabele 4.2 lahko vidimo, da najde druga metoda, ki išče obraze na osnovi parametričnega modela več obrazov kot prva metoda, ki išče obraze na osnovi eksplicitno določenih mej. Izboljšanje se pokaže predvsem pri drugi in tretji bazi, katerih slike so posnete pod nestandardno osvetlitvijo. Metoda izboljša delež pravilno najdenih obrazov tudi pri četrti, peti in šesti bazi, vendar se v zameno zato poveča odstotek nepravilno najdenih obrazov. Čeprav je glavna prednost metode to, da deluje bolje pri nestandardni osvetlitvi lahko vidimo, da je izboljšala uspešnost tudi na slikah prve baze, ki so bile zajete pri normalni osvetlitvi. Prednost metode na osnovi modela je tudi ta, da je metoda prilagodljiva. S pomočjo nastavitve mejne vrednosti lahko sami vplivamo na razmerje med pravilno najdenimi in napačno najdenimi obrazi. Pri preizkusu smo mejno vrednost nastavili tako, da smo dosegali, v primerjavi s prvo metodo, boljše rezultate pri pravilno najdenih obrazih, hkrati pa smo ohranili majhen delež napačno najdenih obrazov.

Slika 4.13 prikazuje primer slike, pri kateri je druga metoda na osnovi parametričnega modela (slika 4.13b) v primerjavi s prvo metodo na osnovi eksplicitno določenih mej (slika 4.13a), uspešneje ohranila kožne barve.

S pomočjo metode na osnovi parametričnega modela smo poskušali izdelati tudi splošen model, na podlagi katerega bomo razločevali kožno barvo različno osvetljenih obrazov (slika 4.14). Slika 4.14 prikazuje različne kožne modele posameznih podatkovnih baz, ki smo jih združili v posplošen primer.

Model smo zgradili tako, da smo združili učne slike vseh podatkovnih baz in na podlagi tako izdelanega modela smo izvedli poskus na slikah posamezne baze ter na vseh slikah iz vseh baz. Rezultate prikazuje tabela 4.3.

Iz tabele 4.3 lahko vidimo, da so rezultati pri splošnem primeru veliko slabši kot v prejšnjem primeru, ko smo preizkušali s prilagojenim modelom. Pri tem preizkusu se tudi izkaže, da prva metoda na osnovi eksplicitno določenih mej deluje boljše. Ugotovimo lahko, da metoda na osnovi parametričnega modela izboljša uspešnost samo na bazi slik, iz katere je bil zgrajen model.

4.2.3 Modeliranje kožne barve različnih kožnih polti

Ker je predlagana metoda na osnovi parametričnega modela zelo prilagodljiva na različne osvetlitve smo hoteli preizkusiti tudi, ali lahko izdelamo model, ki bo uspešno našel tudi

	1. baza		2. baza		3. baza		4. baza		5. baza		6. baza	
	TP	FP	TP	FP	TP	FP	TP	FP	TP	FP	TP	FP
1. metoda	76%	35%	55%	31%	28%	33%	69%	13%	60%	12%	39%	9%
2. metoda	56%	13%	44%	9%	36%	74%	65%	14%	59%	14%	53%	14%
	skupaj											
	TP	FP										
1. metoda	56%	22%										
2. metoda	50%	35%										

Tabela 4.3: Preizkus metode na osnovi posplošenega parametričnega modela nad slikami različnih osvetlitev.

različne kožne polti. Izbrali smo posamezne obraze različnih polti iz sedmih rasnih skupin (slika 4.15):

- azijska rasa,
- azijska rasa srednje Evrope,
- črnska oziroma afroameriška rasa,
- španska rasa,
- ameriška rasa,
- otoško-pacifiška rasa,
- bela rasa.

Model smo zgradili na osnovi majhne učne množice desetih obrazov velikosti 100×100 slikovnih elementov iz podatkovne baze Color FERET [49]. Učni primeri so vsebovali kožno barvo različnih ras. Metodo smo preizkusili na testni sliki 4.15. Metoda na osnovi parametričnega modela je uspešno segmentirala različne kožne barve (slika 4.16). S tem smo pokazali, da je mogoče uspešno izdelati model za določitev kožne barve različnih polti. Da bi lahko preizkusili uspešnost parametrične metode tako izdelanega modela pa nismo imeli ustrezno velike testne baze. Iz enakega razloga tudi nismo mogli zgraditi neparametričnega modela, za katerega potrebujemo veliko učnih primerov.

4.3 Metoda na osnovi neparametričnega modela

Za razliko od parametričnega modela, ki smo ga določili na podlagi parametrov Gaussove porazdelitve, neparametrični model določimo na osnovi barvnega histograma. Kožno barvo smo predstavili s pomočjo kromatičnih vrednosti r in g , in tako določili dvodimenzionalni histogram (slika 4.17).

Na osnovi tako izračunanega histograma kožne barve smo s pomočjo postopka opisanega v poglavju 3.2.2 določili verjetnostno funkcijo P_{skin} :

$$P_{skin} = \frac{skin[c]}{Norm}, \quad (4.7)$$

kjer $skin[c]$ predstavlja število pojavitev barve c v histogramu (višina stolpca na sliki 4.17), konstanto $Norm$ pa smo določili kot maksimalno vrednost v histogramu. Tako izračunana vrednost verjetnosti P_{skin} nam je služila za razločevanje kožne barve.

Mejno vrednost na podlagi katere smo označevali slikovne elemente smo določili po postopku, opisanem v prejšnjem poglavju.

Tudi metodo na osnovi neparametričnega modela smo preizkusili po enakem postopku, kot prej opisano metodo na osnovi parametričnega modela. Razlikoval se je samo postopek učenja modela. Učna množica, s pomočjo katere smo zgradili neparametrične modele je vsebovala približno 100 različnih obrazov. Za ugodno razmerje med velikostjo in zmogljivostjo metode smo se odločili, da bomo kožno barvo predstavili s pomočjo histograma velikosti 128×128 elementov. Rezultate, ki smo jih dobili na osnovi tako zgrajenih neparametričnih modelov prikazuje tabela 4.4.

	1. baza		2. baza		3. baza		4. baza		5. baza		6. baza	
	TP	FP	TP	FP	TP	FP	TP	FP	TP	FP	TP	FP
1. metoda	76%	35%	55%	31%	28%	33%	69%	13%	60%	12%	39%	9%
2. metoda	56%	13%	44%	9%	36%	74%	65%	14%	59%	14%	53%	14%
3. metoda	86%	32%	62%	30%	60%	42%	85%	20%	81%	19%	70%	18%

Tabela 4.4: Preizkus metode na osnovi različnih neparametričnih modelov na slikah različnih osvetlitev.

Vidimo lahko, da deluje metoda na osnovi neparametričnega modela (metoda 3) bolje,

kot metoda na osnovi parametričnega modela (metoda 2) in s tem tudi bolje od prve metode eksplicitno določenih mej (metoda 1). Povečan delež najdenih obrazov pa pripomore tudi k povečanju deleža napačno najdenih obrazov, vendar lahko to pomanjkljivost odpravimo s pomočjo Bayesovega klasifikatorja (poglavje 3.2.2) in dodatnega barvnega modela s katerim predstavimo moteče barve.

Izgradnja neparametričnega modela motečih barv, ki ne predstavljajo kožne barve je enak postopku izgradnje modela kožne barve. Za učne primere izberemo barve ozadja, obleke, predvsem pa tiste barve, ki jih metoda napačno klasificira, da pripadajo kožni barvi. Na osnovi tako zgrajenih modelov kožne barve in modela ostalih motečih barv klasificiramo slikovni element za kožno barvo, če ustreza neenačbi:

$$\frac{P(c|skin)}{P(c|\neg skin)} > \Theta, \quad (4.8)$$

kjer vrednost $P(c|skin)$ označuje verjetnost, da barva c pripada kožni barvi in vrednosti $P(c|\neg skin)$ označuje verjetnosti, da barva c pripada nekožni barvi. Mejno vrednost Θ določimo po enakem postopku, kot smo jo določili pri metodi na osnovi parametričnega modela (poglavje 4.2).

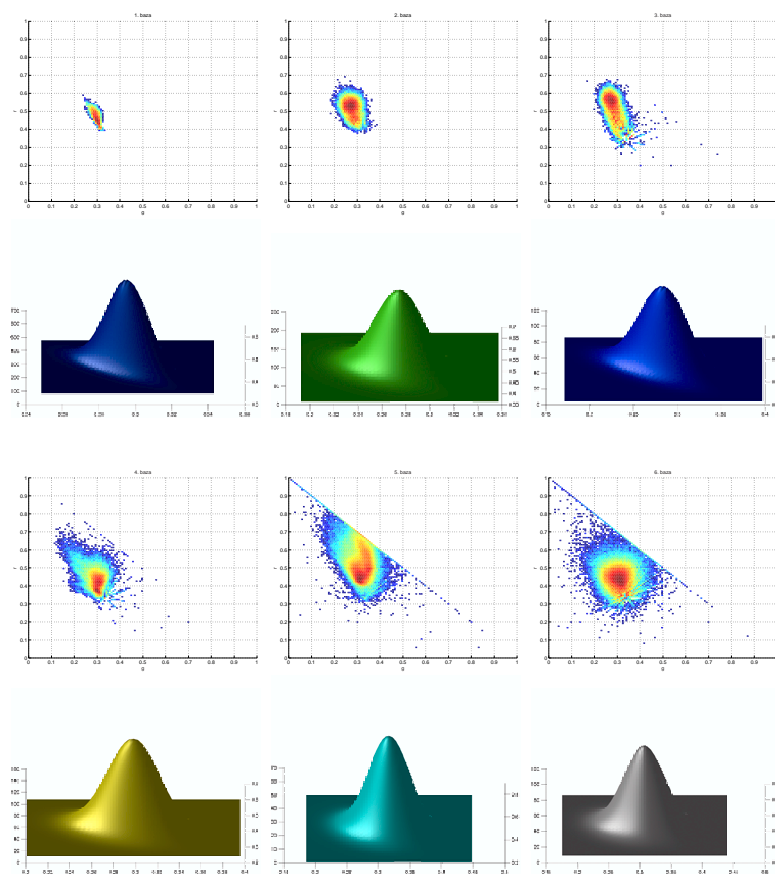
Slika 4.18 prikazuje primer slike, pri kateri nadgrajena neparametrična metoda z modelom nezaželenih barv, izboljša izločanje nekožne barve.

4.4 Zaključek

Po preizkusu vseh treh metod za segmentacijo kožne barve smo ugotovili, da delujeta metodi na osnovi modelov zelo dobro, če sta naučeni nad učno množico primerov enake ali podobne osvetlitve, kot testna množica. Splošni model, zgrajen na osnovi različnih učnih primerov iz vseh podatkovnih baz, pa nam vrača veliko napačno označenih regij.

Metode na osnovi modelov so zelo prilagodljive in so zato zelo primerne v okoljih, kjer lahko prilagajamo model glede na trenutno osvetlitev. Metodo na osnovi parametričnega modela izberemo, ko imamo malo učnih primerov. Ta metoda namreč omogoča interpolacijo. Za boljšo klasifikacijo pa uporabimo metodo na osnovi neparametričnega modela, ki pa ga je težje naučiti. Za učenje namreč potrebujemo veliko učnih primerov, samo učenje pa je tudi časovno in prostorsko zelo zahtevno. Prednost te metode pa je v tem,

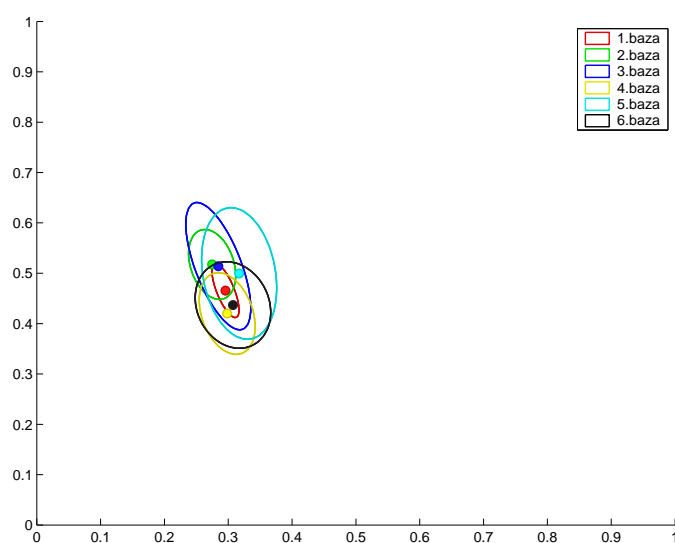
da lahko v kombinaciji z modelom nezaželenih barv močno zmanjšamo število nepravilno označenih regij nekožne barve.



Slika 4.12: Porazdelitev kožne barve in pripadajoči Gaussov model za različne osvatlitve posamezne testne baze v kromatičnem barvnem prostoru. Posamezni pari slik pripadajo posamezni podatkovni bazi. Slika levo zgoraj in slika pod njo prikazujeta porazdelitev in Gaussov model za slike prve baze.



Slika 4.13: Primer ohranjanja kožne barve na testni sliki nestandardne osvetlitve s pomočjo dveh metod: a) metoda eksplicitno določenih mej, b) metoda na osnovi parametričnega modela.



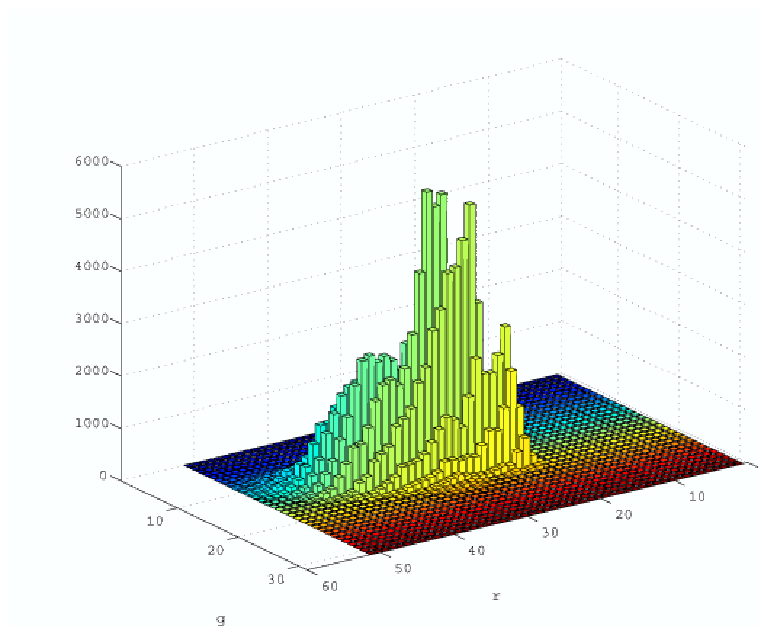
Slika 4.14: Obrisi Gaussovih krivulj modelov kožne barve slik različnih podatkovnih baz.



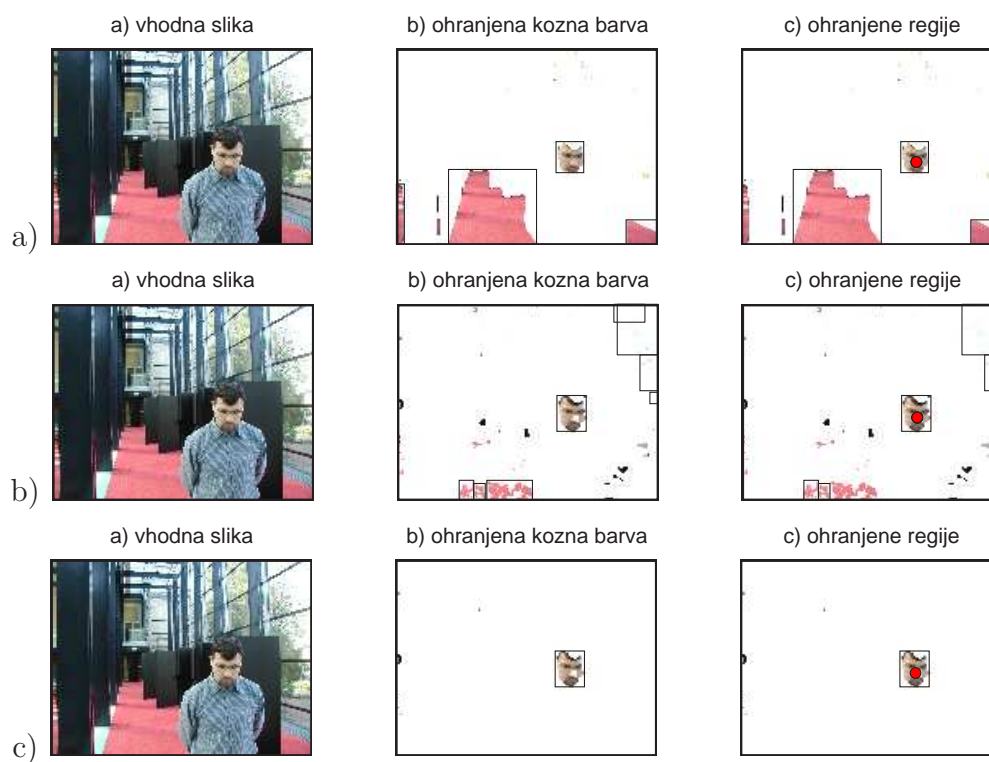
Slika 4.15: Primeri obrazov različnih rasnih skupin (vir: podatkovna baza FERET [49]).



Slika 4.16: Izločene barve obrazov različnih rasnih skupin.



Slika 4.17: Kožna barva predstavljena z dvodimenzionalnim histogramom.



Slika 4.18: Primer slike segmentirane s pomočjo metod: a) eksplicitno določenih mej, b) na osnovi neparametričnega modela, c) na osnovi neparametričnega modela, kateremu smo dodali model neuspešno klasificirane rdeče barve tal.

Poglavje 5

Iskanje obrazov na podlagi vzorcev

Kot smo opisali že v uvodu, je na področju iskanja obrazov veliko metod, ki temeljijo na vzorcih. Takšen pristop je za iskanje obrazov zelo primeren, saj lahko že na pogled opazimo, da so vsi obrazi od spredaj ovalne oblike in pri dovolj majhni ločljivosti opazimo, da so si med seboj zelo podobni - temnejši vodoravni lisi na mestu oči in ust, svetlejša horizontalna lisa na mestu nosu in območja lic ter čela, ki ne vsebujejo veliko teksture. Vzorci, na osnovi katerih preiskujemo sliko, so lahko določeni vnaprej s strani ekspertov, ali pa jih s pomočjo učne množice določimo sami. Na podlagi tega, kako so določeni vzorci, ločimo dve skupini: metode ujemanja s šablono (ang. Template matching) in metode na osnovi videza (ang. Appearance-based methods). Prav slednje so se na področju iskanja obrazov izkazale za zelo uspešne.

Med vodilnimi metodami za iskanje obrazov na sliki veljata metodi na osnovi nevronskih mrež [55] ter na osnovi Naivnega Bayesovega klasifikatorja [59]. V zadnjem času pa se je med metodami pojavila še Bayesova metoda diskriminantnih značilk BDF(ang. Bayesian Discriminating Features) [40], ki se po uspešnosti uvršča v sam vrh metod za iskanju frontalnih obrazov na sivinskih slikah.

V nadaljevanju bomo podrobno predstavili metodo BDF in jo preizkusili na neodvisni množici sivinskih slik obrazov.

5.1 Bayesova metoda diskriminantnih značilk

Bayesova metoda diskriminantnih značilk (ang. Bayesian Discriminating Features) je namenjena iskanju frontalnih obrazov na sivinskih slikah. Metoda obsega tri glavne naloge:

1. analizo diskriminantnih značilk vhodne slike,
2. statistično modeliranje razredov obrazov in neobrazov in
3. uporabo Bayesovega klasifikatorja za iskanje frontalnih obrazov.

Za analizo diskriminantnih značilk uporabimo slikovne elemente vhodne slike, njeno predstavitev s Haarovimi valčki in njeno amplitudno preslikavo. Raziskave na področju Haarovih valčkov so pokazale, da je predstavitev s pomočjo valčkov zelo primerna za predstavitev slik, še posebej pri iskanju obrazov in iskanju pešcev na slikah. Zaradi tega smo tudi sami za enodimenzionalno predstavitev slik uporabil Haarove valčke. Amplitudna preslikava, ki preslika sliko v vertikalni in horizontalni smeri predstavi simetričnost slike v vertikalni smeri, in karakteristike obraza v horizontalni smeri. S takšno kombinacijo vhodne slike, njene Haarove predstavitve in amplitudne preslikave, zgradimo vektor značilk, ki še bolj poudari raznolikost posameznih značilk.

Statistična modela razredov obrazov in ne-obrazov pravzaprav ocenita gostoto pogojne verjetnosti teh dveh razredov. Za razliko od razreda obrazov, ki je običajno modeliran kot normalna porazdelitev več spremenljivk, je razred neobrazov težje modelirati zaradi dejstva, da slike neobrazov predstavljajo vso okolico obraza. Določitev takšne mejne kategorije je v praksi zelo težka. Mi smo podmnožico neobrazov določili tako, da smo izbrali samo tiste slike, ki ležijo zelo blizu razreda obrazov in jih nato modelirali kot normalno porazdelitev več spremenljivk. Idejo za uporabo podmnožice neobrazov pri izdelavi algoritma za iskanje obrazov smo dobili v statistični analizi, ki temelji na metodi podpornih vektorjev (ang. Support Vectors Machine). Pri metodi podpornih vektorjev se pri gradnji sistema upoštevajo samo podporni vektorji oz. vzorci, ki ležijo blizu maksimalnega odmika od hiperravnine. Analogija s to metodo je v tem, da tudi mi izberemo podmnožico neobrazov oz. "podporne neobraz", s pomočjo katerih zgradimo model neobrazov.

Na koncu uporabimo Bayesov klasifikator, s pomočjo katerega iščemo frontalne obraze na sliki. Bayesov klasifikator dopušča minimalno napako pri znani verjetnosti razredov

obrazov in neobrazov. Napaka, imenovana tudi Bayesova napaka, je optimalno merilo za razlikovanje razredov različnih značilk obrazov oziroma neobrazov.

V nadaljevanju bomo natančno opisali posamezne komponente BDF metode: analizo diskriminantnih značilk, izgradnjo modelov obrazov in neobrazov in izgradnjo Bayesovega klasifikatorja za iskanje frontalnih obrazov. Pred tem pa si oglejmo postopek izbire učnih primerov slik obrazov, na podlagi katerih bomo določili značilke, ki jih bomo kasneje uporabili za izgradnjo modela obrazov.

5.2 Normalizacija učnih slik

Na uspešnost metode lahko zelo vpliva izbrana učna množica, zato je pravilna izbira učnih primerov zelo pomembna. Učni primeri pripadajo dvema razredoma, t.j. razredu "obrazov" in razredu "neobrazov". Medtem ko določitev množice obrazov ni težka naloga, je določitev množice neobrazov večji izziv, saj lahko množica neobrazov vsebuje vse slike, ki ne vsebujejo obrazov in teh je neskončno. Zato je za uspešnost metode pomembna tudi pravilna izbira množice neobrazov. Kot smo že omenili v uvodu smo množico neobrazov določili tako, da smo izbrali podmnožico takšnih slik, ki ležijo zelo blizu razreda obrazov.

Za določitev učnih primerov obrazov smo uporabili podatkovno bazo obrazov FERET [49]. Iz podatkovne baze smo izbrali 600 frontalnih obrazov v pokončnem položaju, posnetih pri enakomerni osvetlitvi, velikosti 256×384 slikovnih elementov (slika 5.1). Učne primere za množico neobrazov pa smo izbrali iz desetih naključnih slik narave, na katerih ni bilo obrazov (slika 5.2). Podroben postopek izgradnje učne množice neobrazov je opisan v poglavju 5.6.1.

Učna množica obrazov je vsebovala slike 200 različnih oseb iz množice Batch 15 podatkovne baze FERET. Vsaka oseba je bila posneta trikrat in sicer najprej z običajnim izrazom na obrazu (slike skupine ba*.tif), nato z nasmehom (slike skupine bj*.tif) in končno pri drugačni osvetlitvi (slike skupine bk*.tif) (slika 5.3). Vsi obrazi pa so bili posneti od spredaj in v pokončnem položaju. Nekateri obrazi so vsebovali tudi dodatke, kot so brada ali brki. Nobena oseba iz učne množice pa ni nosila očal.

Same slike portretov ljudi, niso najbolj primerne za učenje, saj poleg obraza vsebujejo še ostale moteče slikovne elemente, kot so obleka, lase ali ozadje. Zaradi tega smo na



Slika 5.1: Primeri obrazov iz podatkovne baze FERET.



Slika 5.2: Primer slike narave, ki ne vsebuje obrazov.

podlagi položaja oči posamezno sliko normalizirali na enotno velikost 16×16 slikovnih elementov. Izbrana velikost 16×16 je bila kasneje tudi velikost našega učnega vzorca. Poleg same slike obraza je podatkovna baza FERET vsebovala še pripadajočo datoteko s koordinatami značilk obraza (pozicijo levega in desnega očesa, pozicijo nosu in pozicijo sredine ustnic). Na podlagi tako določenih koordinat oči smo posamezne slike normalizirali po sledečem postopku (slika 5.4):

1. preberemo koordinate levega in desnega očesa,
2. izračunamo nagib glave in ga ustrezno popravimo,



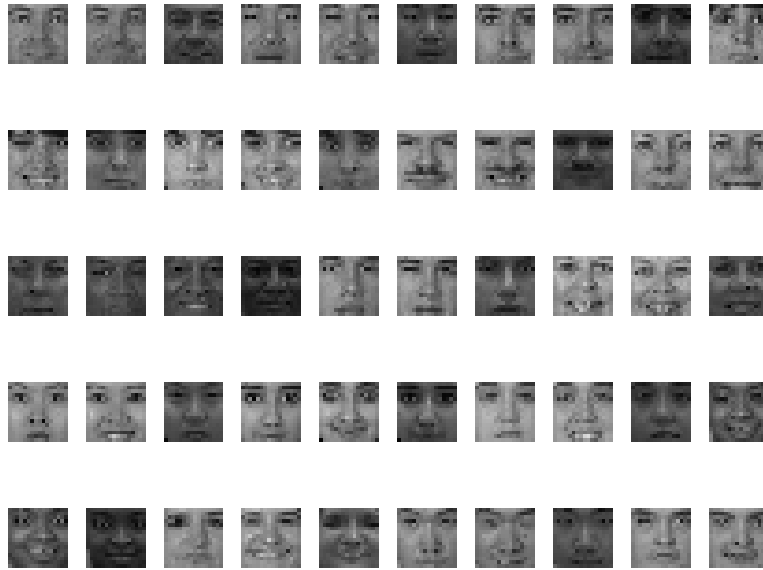
Slika 5.3: Učni primeri istih oseb iz množice obrazov.

3. določimo razdaljo med očmi,
4. na podlagi razdalje sliko ustrezno zmanjšamo,
5. na podlagi pozicije oči iz slike izrežemo samo obraz.



Slika 5.4: Postopek normalizacije slike obraza: a) originalna slika obraza z označenimi koordinatami oči in nagibom glave, b) ustrezno zasukan obraz, c) zmanjšana slika obraza, d) izrezan osrednji del obraza glede na pozicijo oči.

Z opisanim postopkom smo normalizirali vseh 600 obrazov iz podatkovne baze FERET. Normaliziran obraz je bil velikosti 16×16 slikovnih elementov. Nekaj tako normaliziranih obrazov iz učne množice prikazuje slika 5.5.



Slika 5.5: Primeri normaliziranih obrazov učne množice.

5.3 Analiza diskriminantnih značilnk

S pomočjo analize diskriminantnih značilnk zgradimo vektor značilnk, ki močno poveča razliko med značilnkami. Vektor značilnk vsebuje vhodno sliko, njeno enodimenzionalno predstavitev s pomočjo Haarovih valčkov in amplitudno preslikavo slike. Haarovi valčki so se izkazali za zelo učinkovite pri iskanju obrazov in pešcev, s pomočjo amplitudne preslikave pa zajamemo vertikalno simetrijo in horizontalne značilnosti obraza.

Naj matrika $I(i, j) \in \mathbb{R}^{m \times n}$ predstavlja vhodno sliko (t.j. učno sliko razreda obrazov oziroma neobrazov ali posamezen del testne slike) in vektor $\mathbf{X} \in \mathbb{R}^{mn}$ naj vsebuje sliko, kateri smo združili vrstice (ali stolpce) matrike $I(i, j)$. S pomočjo Haarove predstavitve dobimo iz slike $I(i, j)$ dve sliki, $I_h(i, j) \in \mathbb{R}^{(m-1) \times n}$ in $I_v(i, j) \in \mathbb{R}^{m \times (n-1)}$, ki predstavljata horizontalno oziroma vertikalno sliko razlik.

$$I_h(i, j) = I(i+1, j) - I(i, j) \quad 1 \leq i < m, 1 \leq j \leq n \quad (5.1)$$

$$I_v(i, j) = I(i, j+1) - I(i, j) \quad 1 \leq i \leq m, 1 \leq j < n. \quad (5.2)$$

Naj bosta $\mathbf{X}_h \in \mathbb{R}^{(m-1)n}$ in $\mathbf{X}_v \in \mathbb{R}^{m(n-1)}$ vektorja, ki združujeta posamezne vrstice

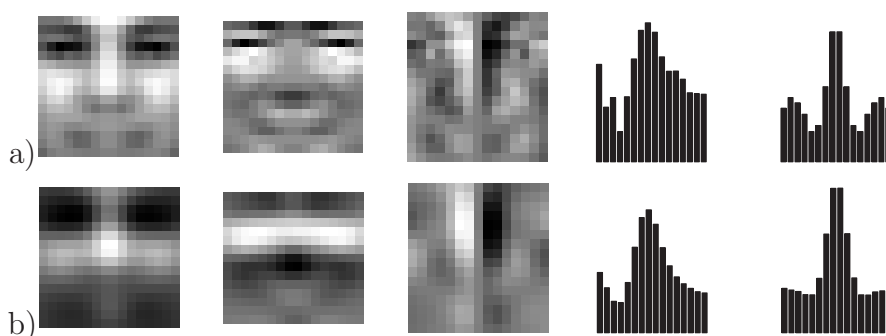
(ali stolpce) matrike $I_h(i, j)$ in matrike $I_v(i, j)$.

Amplitudna preslikava slike $I(i, j)$ po vrsticah in po stolpcih oblikuje horizontalno in vertikalno predstavitev slike, $\mathbf{X}_r \in \mathbb{R}^m$ in $\mathbf{X}_c \in \mathbb{R}^n$.

$$\mathbf{X}_r(i) = \sum_{j=1}^n I(i, j) \quad 1 \leq i \leq m \quad (5.3)$$

$$\mathbf{X}_c(j) = \sum_{i=1}^m I(i, j) \quad 1 \leq j \leq n. \quad (5.4)$$

Postopek diskriminantne analize značiln na povprečni sliki obraza in povprečni sliki neobraza prikazuje slika 5.6. Prva slika predstavlja povprečni obraz, druga in tretja slika sta Haarova predstavitev slike obraza, zadnja dva grafa pa prikazujeta njeno amplitudo. Vidimo lahko da sta predstavitvi slik obraza in neobraza, ter njuni amplitudi zelo podobni. Vzrok je izbira učnih slik neobrazov, ki ležijo zelo blizu razreda obrazov.



Slika 5.6: Diskriminantna analiza značiln povprečnega obraza in neobraza.

Preden zgradimo vektor značiln, normaliziramo vektorje \mathbf{X} , \mathbf{X}_h , \mathbf{X}_v , \mathbf{X}_r in \mathbf{X}_c tako, da jim odštejemo povprečno vrednost njihovih vrednosti in delimo z njihovimi standardnimi deviacijami. Tako normalizirane vektorje označimo z, $\tilde{\mathbf{X}}$, $\tilde{\mathbf{X}}_h$, $\tilde{\mathbf{X}}_v$, $\tilde{\mathbf{X}}_r$ in $\tilde{\mathbf{X}}_c$. Novi vektor značiln $\tilde{\mathbf{Y}} \in \mathbb{R}^N$ je sedaj definiran kot skupek normaliziranih vektorjev:

$$\tilde{\mathbf{Y}} = (\tilde{\mathbf{X}}^t \tilde{\mathbf{X}}_h^t \tilde{\mathbf{X}}_v^t \tilde{\mathbf{X}}_r^t \tilde{\mathbf{X}}_c^t)^t, \quad (5.5)$$

kjer je t oznaka za transponiranje in $N = 3mn$ je dimenzija vektorja značilk $\tilde{\mathbf{Y}}$. Končni vektor značilk $\mathbf{Y} \in \mathbb{R}^N$ je normaliziran vektor $\tilde{\mathbf{Y}}$. Tako normaliziran vektor \mathbf{Y} predstavlja končni vektor značilk, ki združuje vhodno sliko, njeno enodimenzionalno Haarovo predstavitev in amplitudno preslikavo:

$$\mathbf{Y} = \frac{\tilde{\mathbf{Y}} - \mu}{\sigma}, \quad (5.6)$$

kjer sta μ in σ povprečna vrednost in standardna deviacija posameznih komponent vektorja $\tilde{\mathbf{Y}}$.

5.4 Statistična modela obrazov in neobrazov

Glavni cilj statističnega modeliranja obrazov in neobrazov je ocena funkcije pogojne gostote verjetnosti PDF (ang. conditional Probability Density Function) teh dveh razredov. Medtem, ko vsebuje razred obrazov samo obraze, vsebuje razred neobrazov vse ostalo, oziroma vse preostale slike. Razumljivo je predpostaviti, da tvori razred obrazov večrazsežno normalno porazdelitev, kar pa ne moremo trditi za razred neobrazov. Razred neobrazov izberemo iz podmnožice slik, ki so blizu samemu modelu obrazov in jih tudi modeliramo kot večrazsežno normalno porazdelitev. Takšna rešitev je podobna izbiri podpornega vektorja pri metodi podpornih vektorjev SVM (ang. Support Vector Machines). Podporni vektorji so tisti učni primeri, ki so najbližji odločitveni hiperravnini SVM-ja in so zato tudi najpomembnejši podatki pri določitvi optimalne pozicije odločitvene hiperravnine. Enako idejo smo uporabili pri naši metodi, ko smo izbrali "podporne neobraze", ki ležijo najbližje razredu obrazov. S tem smo določili optimalno odločitveno ravnino med obrazi in neobrazi.

5.4.1 Model obraza

Funkcija pogojne verjetnosti razreda obrazov, ω_f je modelirana kot večrazsežna normalna porazdelitev:

$$p(\mathbf{Y}|\omega_f) = \frac{1}{(2\pi)^{N/2}|\Sigma_f|^{1/2}} \exp \left\{ -\frac{1}{2}(\mathbf{Y} - \mathbf{M}_f)^t \Sigma_f^{-1} (\mathbf{Y} - \mathbf{M}_f) \right\}, \quad (5.7)$$

kjer sta $\mathbf{M}_f \in \mathbb{R}^N$ in $\Sigma_f \in \mathbb{R}^{N \times N}$ povprečna vrednost in kovariančna matrika razreda obrazov ω_f . Če obe strani pomnožimo z naravnim logaritmom dobimo:

$$\ln [p(\mathbf{Y}|\omega_f)] = -\frac{1}{2} \{ (\mathbf{Y} - \mathbf{M}_f)^t \Sigma_f^{-1} (\mathbf{Y} - \mathbf{M}_f) + N \ln(2\pi) + \ln |\Sigma_f| \}. \quad (5.8)$$

Kovariančno matriko Σ_f , lahko razčlenimo s pomočjo analize glavnih komponent PCA (ang. Principal Component Analysis) v sledečo obliko:

$$\Sigma_f = \Phi_f \Lambda_f \Phi_f^t, \text{ kjer velja: } \Phi_f \Phi_f^t = \Phi_f^t \Phi_f = \mathbf{I}_N, \quad \Lambda_f = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_N\}, \quad (5.9)$$

V enačbi 5.9 imamo matriko ortogonalnih lastnih vektorjev $\Phi_f \in \mathbb{R}^{N \times N}$ in na diagonalni matrike $\Lambda_f \in \mathbb{R}^{N \times N}$ lastne vrednosti v padajočem vrstnem redu ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$). Matrika \mathbf{I}_N je identična matrika ali matrika identitete. Pomembna lastnost analize glavnih komponent je optimalna konstrukcija signala v smislu najmanjše povprečne kvadratne napake, kjer je za predstavitev glavnega signala uporabljena samo podmnožica glavnih komponent. Glavne komponente so predstavljene z vektorjem $\mathbf{Z} \in \mathbb{R}^N$:

$$\mathbf{Z} = \Phi_f^t (\mathbf{Y} - \mathbf{M}_f). \quad (5.10)$$

Iz enačb 5.8, 5.9, 5.10 nato sledi:

$$\ln [p(\mathbf{Y}|\omega_f)] = -\frac{1}{2} \{ \mathbf{Z}^t \Lambda_f^{-1} \mathbf{Z} + N \ln(2\pi) + \ln |\Lambda_f| \}. \quad (5.11)$$

Posamezne vrednosti vektorja \mathbf{Z} so glavne komponente. Z uporabo lastnosti optimalne konstrukcije signala pri analizi glavnih komponent [46] uporabimo samo prvih M ($M \ll N$) glavnih komponent, s pomočjo katerih določimo funkcijo pogojne verjetnosti. Nadalje uporabimo model, ki upošteva preostale komponente $N - M$ lastnih vrednosti, $\lambda_{M+1}, \lambda_{M+2}, \dots, \lambda_N$ s povprečenjem teh vrednosti:

$$\rho = \frac{1}{N - M} \sum_{k=M+1}^N \lambda_k. \quad (5.12)$$

V enačbi 5.10 imamo $\|\mathbf{Z}\|^2 = \|\mathbf{Y} - \mathbf{M}_f\|^2$, kjer z oznako $\|\cdot\|$ označujemo normalno normo. Ta enačba nam dokazuje, da transformacija s pomočjo analize glavnih komponent ne spremeni norme. Sedaj iz enačb 5.11 in 5.12 sledi:

$$\ln [p(\mathbf{Y}|\omega_f)] = -\frac{1}{2} \left\{ \sum_{i=1}^M \frac{z_i^2}{\lambda_i} + \frac{\|Y - M_f\|^2 - \sum_{i=1}^M z_i^2}{\rho} + \ln \left(\prod_{i=1}^M \lambda_i \right) + (N - M) \ln \rho + N \ln(2\pi) \right\}, \quad (5.13)$$

kjer so vednosti z_i posamezne komponente vektorja \mathbf{Z} , ki smo ga definirali z enačbo 5.10. Iz enačbe 5.13 vidimo, da lahko funkcijo pogojne verjetnosti PDF določimo s prvimi M vrednostmi glavnih komponent vektorja \mathbf{Z} , z vhodno sliko \mathbf{Y} , s povprečno vrednostjo obraza \mathbf{M}_f in z lastnimi vrednostmi razreda obrazov λ_i .

5.4.2 Model neobraza

Modeliranje razreda neobrazov se začne z generiranjem vzorcev neobrazov iz slik, ki ne vsebujejo obrazov. Za vsak vzorec, ki ga izrežemo iz večje slike, izračunamo vrednost funkcije iz enačbe 5.13. Na podlagi izračunane vrednosti izberemo samo tisto podmnožico slik ali vzorcev, ki ležijo blizu razreda obrazov. Tako izbrana podmnožica je nato naša učna množica neobrazov, s pomočjo katere določimo funkcijo pogojne verjetnosti za razred neobrazov ω_n , ki jo ravno tako kot množico obrazov modeliramo kot večrazsežno normalno porazdelitev:

$$p(\mathbf{Y}|\omega_n) = \frac{1}{(2\pi)^{N/2} |\Sigma_n|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{Y} - \mathbf{M}_n)^t \Sigma_n^{-1} (\mathbf{Y} - \mathbf{M}_n) \right\}, \quad (5.14)$$

kjer sta $\mathbf{M}_n \in \mathbb{R}^N$ in $\Sigma_n \in \mathbb{R}^{N \times N}$ povprečna vrednost in kovariančna matrika razreda neobrazov ω_n .

Razčlenitev kovariančne matrike Σ_n s pomočjo analize glavnih komponent nam da:

$$\Sigma_n = \Phi_n \Lambda_n \Phi_n^t, \text{ kjer velja: } \Phi_n \Phi_n^t = \Phi_n^t \Phi_n = \mathbf{I}_N, \Lambda_n = \text{diag}\{\lambda_1^{(n)}, \lambda_2^{(n)}, \dots, \lambda_N^{(n)}\}, \quad (5.15)$$

kjer je $\Phi_n \in \mathbb{R}^{N \times N}$ matrika ortogonalnih lastnih vektorjev, $\Lambda_n \in \mathbb{R}^{N \times N}$ diagonalna matrika lastnih vrednosti v padajočem vrstnem redu ($\lambda_1^{(n)} \geq \lambda_2^{(n)} \geq \dots \geq \lambda_N^{(n)}$) in $\mathbf{I}_N \in \mathbb{R}^{N \times N}$ identična matrika. Vektor glavnih komponent $\mathbf{U} \in \mathbb{R}^N$ je definiran kot:

$$\mathbf{U} = \Phi_n^t(\mathbf{Y} - \mathbf{M}_n). \quad (5.16)$$

Preostanek vrednosti glavnih komponent $N - M$, $\lambda_{M+1}^{(n)}, \lambda_{M+2}^{(n)}, \dots, \lambda_N^{(n)}$ določimo s povprečjem:

$$\varepsilon = \frac{1}{N - M} \sum_{k=M+1}^N \lambda_k^{(n)}. \quad (5.17)$$

Funkcija pogojne verjetnosti PDF za razred neobrazov je tako določena z enačbo:

$$\ln [p(\mathbf{Y}|\omega_n)] = -\frac{1}{2} \left\{ \sum_{i=1}^M \frac{u_i^2}{\lambda_i^{(n)}} + \frac{\|\mathbf{Y} - \mathbf{M}_n\|^2 - \sum_{i=1}^M u_i^2}{\varepsilon} + \ln \left(\prod_{i=1}^M \lambda_i^{(n)} \right) + (N - M) \ln \varepsilon + N \ln(2\pi) \right\}, \quad (5.18)$$

kjer so vrednosti u_i posamezne komponente vektorja \mathbf{U} , ki smo ga definirali z enačbo 5.16. Iz enačbe 5.18 vidimo, da lahko funkcijo pogojne verjetnosti PDF določimo s prvimi M vrednostmi glavnih komponent vektorja \mathbf{U} , z vhodno sliko \mathbf{Y} , s povprečno vrednostjo neobrazov \mathbf{M}_n in z lastnimi vrednostmi razreda obrazov $\lambda_i^{(n)}$.

5.5 Bayesov klasifikator za iskanje frontalnih obrazov

Po izgradnji statističnih modelov obrazov in neobrazov, uporabimo za iskanje obrazov na sliki Bayesov klasifikator. Značilnost Bayesovega klasifikatorja je v tem, da pri znanih vrednostih porazdelitvenih funkcij razredov obrazov in neobrazov minimizira napako. Napaka, imenovana tudi Bayesova napaka, predstavlja optimalno merilo za klasifikacijo diskriminantnih značilk, medtem ko preverja pripadnost posameznemu razredu.

Naj bo $\mathbf{Y} \in \mathbb{R}^N$ vektor značilk, ki predstavlja vhodni vzorec, izrezan iz testne slike. Naj bodo aposteriorne (ang. posteriori) verjetnosti razreda obrazov (ω_f) in razreda neobrazov (ω_n) pri znanem vektorju značilk \mathbf{Y} , $P(\omega_f|\mathbf{Y})$ in $P(\omega_n|\mathbf{Y})$. Vhodni vzorec se klasificira v razred obrazov ali v razred neobrazov glede na Bayesovo odločitveno pravilo:

$$\mathbf{Y} \in \begin{cases} \omega_f & \text{če je } P(\omega_f|\mathbf{Y}) > P(\omega_n|\mathbf{Y}) \\ \omega_n & \text{sicer.} \end{cases} \quad (5.19)$$

Bayesovo pravilo optimizira ločljivost med razredoma v smislu Bayesove napake, kar pripomore k določitvi najboljšega klasifikatorja za iskanje obrazov.

S pomočjo Bayesovega teorema lahko določimo aposteriorni verjetnosti $P(\omega_f|\mathbf{Y})$ in $P(\omega_n|\mathbf{Y})$ tako, da izračunamo pogojni verjetnostni $p(\mathbf{Y}|\omega_f)$ in $p(\mathbf{Y}|\omega_n)$ s pomočjo enačb, ki smo jih predstavili v poglavju 5.4.1 in 5.4.2:

$$P(\omega_f|\mathbf{Y}) = \frac{P(\omega_f)p(\mathbf{Y}|\omega_f)}{p(\mathbf{Y})}, \quad P(\omega_n|\mathbf{Y}) = \frac{P(\omega_n)p(\mathbf{Y}|\omega_n)}{p(\mathbf{Y})}, \quad (5.20)$$

kjer sta $P(\omega_f)$ in $P(\omega_n)$ apriorni verjetnosti razredov obrazov (ω_f) in neobrazov (ω_n) in $P(\mathbf{Y})$ porazdeljena verjetnost vektorja značilnk.

Iz enačb 5.13, 5.18 in 5.20 je Bayesovo pravilo za iskanje obrazov definirano takole:

$$\mathbf{Y} \in \begin{cases} \omega_f & \text{če je } \delta_f + \tau < \delta_n \\ \omega_n & \text{sicer,} \end{cases} \quad (5.21)$$

kjer so posamezni parametri δ_f , δ_n in τ definirani takole:

$$\delta_f = \sum_{i=1}^M \frac{z_i^2}{\lambda_i} + \frac{\|\mathbf{Y} - \mathbf{M}_f\|^2 - \sum_{i=1}^M z_i^2}{\rho} + \ln \left(\prod_{i=1}^M \lambda_i \right) + (N - M) \ln \rho \quad (5.22)$$

$$\delta_n = \sum_{i=1}^M \frac{u_i^2}{\lambda_i^{(n)}} + \frac{\|\mathbf{Y} - \mathbf{M}_n\|^2 - \sum_{i=1}^M u_i^2}{\varepsilon} + \ln \left(\prod_{i=1}^M \lambda_i^{(n)} \right) + (N - M) \ln \varepsilon \quad (5.23)$$

$$\tau = 2 \ln \left[\frac{P(\omega_n)}{P(\omega_f)} \right]. \quad (5.24)$$

Parametra δ_f in δ_n lahko izračunamo s pomočjo vhodnega vzorca \mathbf{Y} , parametrov razreda (povprečna vrednost obraza, prvih M lastnih vektorjev in prvih M lastnih vrednosti). Parameter τ pa je konstanta, ki služi kot kontrolni parameter: večja kot je vrednost, manj je napačno najdenih obrazov (ang. false detection). Za boljšo kontrolo napačno

najdenih obrazov uporabljamo dodaten parameter θ , ki dodatno omeji razreda obrazov in neobrazov. Z upoštevanjem obeh kontrolnih parametrov dobimo sledeče pravilo:

$$\mathbf{Y} \in \begin{cases} \omega_f & \text{če je } (\delta_f < \theta) \text{ in } (\delta_f + \tau < \delta_n) \\ \omega_n & \text{sicer,} \end{cases} \quad (5.25)$$

Kontrolna parametra τ in θ določimo na podlagi preizkusa na množici primerov neodvisnih od testne množice.

5.6 Preizkus metode

Učno množico metode BDF smo zgradili s pomočjo 600 slik frontalnih obrazov podatkovne baze FERET [49] in 10 panoramskih slik narave. Učno množico obrazov smo razširili še s simetrično preslikanimi slikami obrazov in tako dobili 1200 učnih primerov. Na osnovi takšne učne množice smo nato zgradili model obrazov. Iz 10 panoramskih slik pa smo izrezali 4500 primerov slik neobrazov, ki so ležali zelo blizu razredu obrazov. Na tako izbrani učni množici neobrazov smo zgradili model neobrazov. Za preizkus metode BDF pa smo uporabili tri različne testne množice: SET1, SET2 in SET3. Množica SET1 je vsebovala slike frontalnih obrazov iz podmnožic slik Batch 12, 13 in 14 iz podatkovne baze FERET. Slike so večinoma vsebovale samo obraz ali pa obraz in ramena (slika 5.7). Množica SET2 je vsebovala slike frontalnih obrazov množice Batch 2 iz podatkovne baze FERET. Baza je večinoma vsebovala posnetke oseb do pasu (slika 5.8). Glavna razlika med množicama SET1 in SET2 je tudi v tem, da množica SET2 vsebuje tudi osebe z očali in pri nekaterih med njimi se pojavi tudi odboj svetlobe.

Testna baza SET3 pa je vsebovala slike, ki so vsebovale frontalne obraze iz podatkovne baze MIT-CMU [55, 66]. Množici SET1 in SET2, ki vsebujeta 511 in 296 slik vsebujeta tudi enako število obrazov, saj vsaka slika vsebuje samo en obraz. Medtem ko vsebuje množica SET3 80 slik s skupno 227 obrazi (tabela 5.1).

Ker išče metoda BDF samo pokončne frontalne človeške obraze, v testno množico SET3 nismo dodali slik, ki so vsebovala močno zasukane obraze, narisane obraze, obraze risanih junakov, obraze figuric igralnih kart in zakrite obraze. Za razliko od testnih množic SET1 in SET2, ki smo ju sestavili iz iste podatkovne baze FERET, združuje množica

baza	Izvor	št. slik	št. obrazov
SET1	FERET (Batch 12, 13 in 14)	511	511
SET2	FERET (Batch 2)	296	296
SET3	MIT-CMU testna množica	80	227

Tabela 5.1: Testne baze: SET1, SET2 in SET3.

SET3 slike iz različnih virov. Množica SET3 vsebuje slike iz interneta, preslikane slike iz kakovostnih fotografij ter iz časopisnih člankov ter slike zajete z videa. Slike vsebujejo tudi različno velike obraze (slika 5.9), različno nagnjene obraze (slika 5.10), zelo velike obraze (slika 5.11) ali pa zelo majhne obraze (slika 5.12), obraze na slikah slabe kvalitete (slika 5.13) in delno zakrite ali rahlo obrnjene obraze (slika 5.14). V nadaljevanju bomo predstavili postopek statističnega učenja metode BDF ter preizkus na vseh treh testnih bazah slik.

5.6.1 Statistično učenje metode BDF

Za izgradnjo statističnega modela obrazov in neobrazov, moramo s pomočjo učnih primerov določiti pripadajoče parametre teh dveh razredov. Parametre razreda obrazov izračunamo po sledečem postopku:

1. Normaliziramo 600 slik iz podatkovne baze FERET [49] (poglavje 5.2).
2. Učno množico dopolnimo z zrcalnimi slikami učnih primerov in s tem povečamo število učnih primerov na 1200.
3. Izračunamo enodimenzionalno Haarovo predstavitev slike in njeno amplitudno projekcijo in skupaj z vhodno sliko sestavimo vektor značilk, kot je podrobno opisano v poglavju 5.3. Slika 5.6a prikazuje povprečno sliko obraza in njeno Haarovo ter amplitudno predstavitev.
4. Izračunamo parametre razreda, kot so povprečna vrednost obraza, lastni vektorji razreda in lastne vrednosti (poglavje 5.4.1).

5. Določimo parameter M , ki določa število glavnih komponent, ki jih uporabimo za modeliranje verjetnostne funkcije PDF razreda obrazov. Parameter M določa kompromis med uspešnostjo metode in kompleksnostjo izračuna. Za preizkus metode smo uporabili vrednost 10.

Učenje razreda neobrazov pa se začne z zbiranjem učnih primerov neobrazov iz desetih panoramskih slik narave (slika 5.15), ki ne vsebujejo nobenega obraza. Na posameznemu delu slike, standardne velikosti 16×16 slikovnih elementov izračunamo verjetnostno funkcijo PDF (enačba 5.13) razreda obrazov in med učne primere neobrazov dodamo samo tiste slike, ki ležijo zelo blizu razreda obrazov (funkcija PDF ima najnižjo vrednost). S pomočjo tega postopka smo na podlagi panoramskih slik zbrali 4500 učnih primerov neobrazov. Slika 5.6b prikazuje povprečno sliko neobraza in njeno Haarovo ter amplitudno predstavitev. Zaradi dejstva, da so slike neobrazov izbrane tako, da ležijo zelo blizu razreda obrazov lahko vidimo, da je predstavitev neobraza zelo podobna predstavitvi slike obraza. Parametre razreda neobrazov izračunamo po enakem postopku, kot smo to storili za razred obrazov.

Na koncu moramo določiti še kontrolna parametra τ in θ , ki določata število napačnih detekcij. Parametra smo izbrali na podlagi učnih slik in sicer smo za parameter τ uporabili vrednost 100 in za parameter θ vrednost 600.

5.6.2 Rezultati preizkusa

Metodo BDF smo preizkusili na treh testnih bazah: SET1, SET2 in SET3. Medtem, ko sta bazi SET1 in SET2 iz podatkovne baze FERET vsebovali samo en obraz na sliki, je bila testna množica SET3 kompleksnejša in je vsebovala več obrazov. Tabela 5.2 prikazuje uspešnost iskanja frontalnih obrazov na posameznih testnih množicah.

Vidimo lahko, da je metoda brez napačnih detekcij uspešno našla 507 obrazov od 511 iz množice SET1. Slika 5.7 prikazuje primere slik iz testne množice SET1, kjer je najden obraz označen z belim kvadratom. Velikost posamezne testne slike je 256×384 slikovnih elementov, iskanje pa je potekalo pri različnih velikostih. Posamezno sliko smo zmanjšali za faktor $16/300$ in jo postopoma za korak 5 povečevali do faktorja $16/130$. Število x v imenovalcu ulomka $16/x$ predstavlja tudi velikost obraza, ki ga metoda dobi pri tako

baza	št. slik	št. obrazov	št. pravilno najdenih (TP)	št. napačno najdenih (FP)
SET1	511	511	507	0
SET2	296	296	290	0
SET3	80	227	221	1
Skupaj	887	1034	1018	1

Tabela 5.2: Rezultati preizkusa na testnih bazah: SET1, SET2 in SET3.

izbranim faktorju. Na primer, če metoda dobi obraz pri vrednosti imenovalca 16, to pomeni, da smo na originalni sliki ($16/16 = 1$) našli obraz velikosti 16×16 , kar ravno ustreza velikosti vzorca. V množici SET2 je metoda uspešno našla 290 obrazov od 296, ponovno brez napačnih detekcij. Na tej množici smo slike preiskovali od faktor 16/80 do faktorja 16/150 ustrezno zmanjšane slike. Slika 5.8 prikazuje primere slik iz množice SET2, med katerimi so tudi osebe z očali, pri katerih je viden svetlobni odsev. Tretjo množico SET3 pa so sestavljali testni primeri slik podatkovne baze MIT-CMU [55, 66]. Množica je vsebovala 80 različnih slik, ki so skupaj vsebovale 227 obrazov. Slike 5.9, 5.10, 5.11 5.12 prikazujejo nekaj različnih primerov slik iz te testne množice. Ker so testne slike množice SET3 vsebovale različno velike, ter različno nagnjene obraze, smo vsako sliko preiskovali pri različnih velikostih, ter pri različnih nagibih. Tako smo za vsako sliko zgradili piramido dvajsetih različnih velikosti od faktorja 16/20 do 16/55, kjer smo imenovalce povečevali za korak 2. Da smo našli tudi nagnjene obraze smo sliko zarotirali za določen kot, $\pm 5^\circ$, $\pm 10^\circ$, $\pm 15^\circ$ in $\pm 20^\circ$. Metoda BDF se je izkazala za zelo uspešno tudi na slabo kvalitetnih slikah (slika 5.13) ter na slikah, kjer je bil obraz delno zakrit (slika 5.14).

Metoda ni našla šestih obrazov iz množice SET3. Slika 5.16 prikazuje primer slike, pri kateri metoda ni našla skrajno levega obraza. Na vseh slikah množice SET3 je metoda uspešno našla 221 obrazov od 227. Napačno je našla en obraz, ki je prav tako prikazan na sliki 5.16. Če združimo vse rezultate nad vsemi testnimi množicami s skupno 1034 obrazi, lahko ugotovimo, da je uspešnost metode BDF 98.5%, z enim napačno najdenim obrazom.

5.6.3 Primerjava uspešnosti metode

Metodo BDF smo primerjali s trenutno najuspešnejšo metodo na osnovi naivnega Bayesovega klasifikatorja, ki sta jo razvila Schneiderman in Kanade [59]. Metoda je tudi javno dostopna na internetnem naslovu : <http://www.vasc.ri.cmu.edu/cgi-bin/demos/findface.cgi>. Metoda vsebuje dve mejni vrednosti, ki določata razmerje med uspešno najdenimi in napačno najdenimi obrazy. Če povečujemo mejni vrednosti, se zmanjšujeta tako uspešnost kot tudi število napačno najdenih obrazov. Tabela 5.3 prikazuje primerjavo rezultatov med metodo na osnovi naivnega Bayesovega klasifikatorja pri različnih mejnih vrednostih in metodo BDF. Mejne vrednosti metode na osnovi naivnega Bayesovega klasifikatorja so pripisane v oklepajih. Obe metodi smo preizkusili na testni množici SET3, ki je vsebovala slike frontalnih obrazov iz podatkovnih baz MIT-CMU [55, 66].

metoda	št. pravilno najdenih (TP)	št. napačno najdenih (FP)	uspešnost
naivni Bayes. klas. (1.0, 1.0)	218	41	96.0%
naivni Bayes. klas. (2.0, 2.0)	214	5	94.3%
naivni Bayes. klas. (3.0, 3.0)	208	1	91.6%
BDF	221	1	97.4%

Tabela 5.3: Primerjava dveh metod na testni množici SET3 z 227 obrazy.

Iz rezultatov vidimo, da je metoda BDF uspešnejša od metode naivnega Bayesovega klasifikatorja. Vedeti pa moramo, da išče metoda BDF samo frontalne obraze, medtem ko išče druga metoda tudi obraze s profila.

5.7 Zaključek

Na osnovi učne množice smo zgradili modela obrazov in neobrazov, ter Bayesov klasifikator, ki išče frontalne obraze na sivinski sliki [40]. Uspešnost metode smo preizkusili na testnih množicah SET1, SET2 in SET3 in jo primerjali s trenutno najuspešnejšo metodo na osnovi naivnega Bayesovega klasifikatorja, avtorjev Schneiderman in Kanade. Rezultati so pokazali, da metoda BDF uspešno najde 98.5% vseh frontalnih obrazov na sliki.

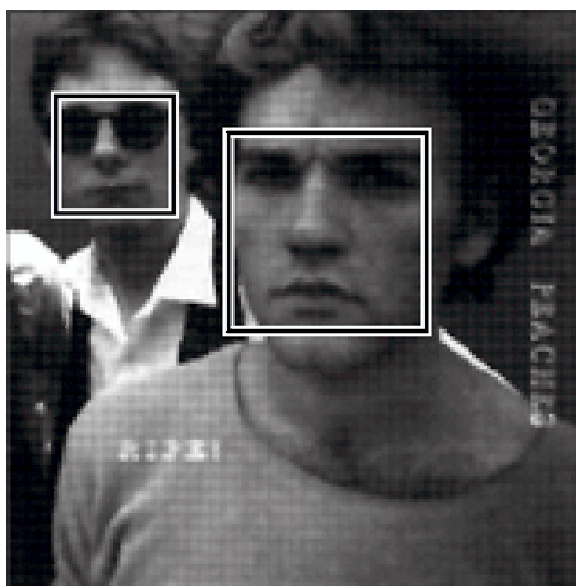
Glavni slabosti metode sta, da najde na sliki samo frontalne obraze, in njena časovna zahtevnost. Za sliko velikosti 320×240 potrebuje metoda v prototipni izvedbi, s pomočjo programskega orodja Matlab, namreč 112 sekund na 1.8 GHz računalniku Pentium 4. Optimiziran program na 900 MHz Sunovi postaji pa potrebuje za to isto sliko 1 sekundo [40]. Kljub optimizaciji pa je metoda zelo počasna, če upoštevamo, da moramo sliko preiskovati pri različnih velikostih in pri različno zasukanih slikah. Zaradi tega smo se odločili združiti metodo na osnovi barve in metodo BDF, kar bomo predstavili v naslednjem poglavju.



Slika 5.7: Primeri najdenih obrazov iz testne baze SET1.



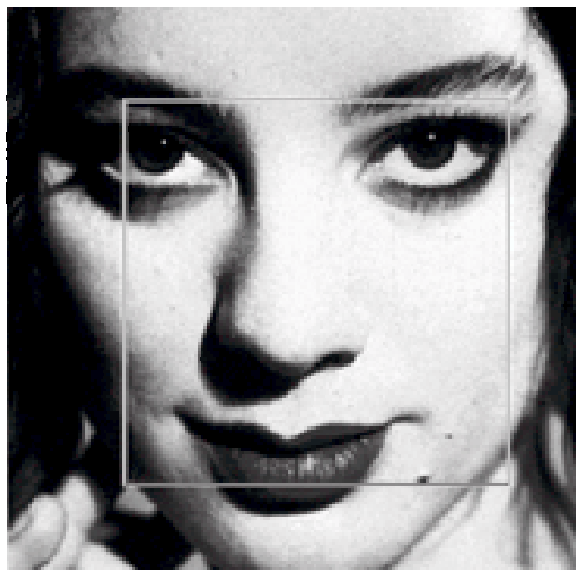
Slika 5.8: Primeri najdenih obrazov iz testne baze SET2.



Slika 5.9: Primer iskanja različno velikih obrazov.



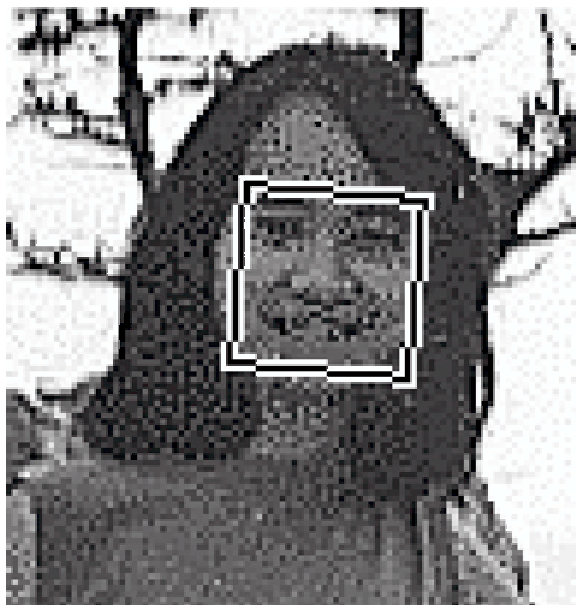
Slika 5.10: Primer iskanja različno nagnjenih obrazov.



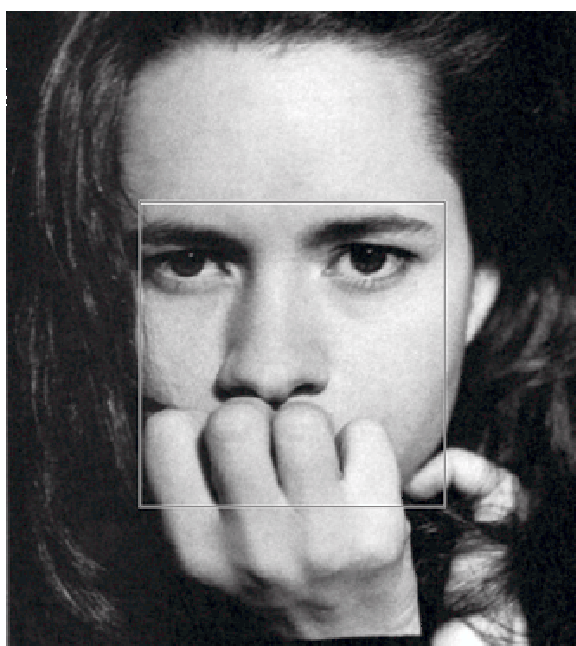
Slika 5.11: Primer uspešno najdenega velikega obraza.



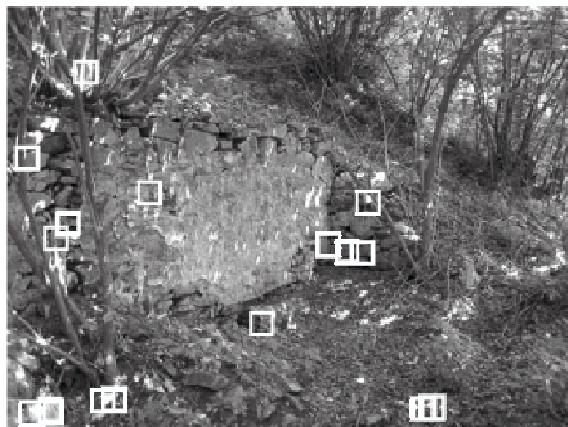
Slika 5.12: Primer uspešno najdenega majhnega obraza.



Slika 5.13: Primer iskanja obrazov na sliki slabe kvalitete.



Slika 5.14: Primer delno zakritega obraza.



Slika 5.15: Izbrani učni primeri neobrazov na sliki iz narave.



Slika 5.16: Primer neuspešno najdenega obraza ter napačno najdenega obraza.

Poglavje 6

Združitev metode na osnovi barve in metode BDF

V tem poglavju bom predstavili novo metodo, ki bo združevala metodo na osnovi barv (poglavje 4) in metodo BDF (poglavje 5). S pomočjo tako izdelane metode bom poskušal doseči večjo hitrost in večjo uspešnost pri iskanju frontalnih obrazov na barvnih slikah.

Metode na osnovi barv so zelo hitre metode. Glavna slabost teh metod pa je v tem, da najdejo poleg obrazov tudi druge dele telesa, ki so tudi kožne barve. Število napačno najdenih obrazov lahko zmanjšamo tako, da izbrane regije, ki predstavljajo kandidate za obraz, preverimo s preprostimi pravili, na podlagi katerih zavrremo izbrane regije, ki so premajhne, prevelike ali nepravilnih razmerij (poglavje 4.2.1). Uporaba takšnih hevrističnih pravil zelo zmanjša število napačno najdenih obrazov, vendar še vedno ne dovolj (tabela 6.1).

Naša metoda uporablja za potrjevanje obrazov metodo BDF, ki uspešno najde frontalne

metoda na osnovi barv	št. slik	št. obrazov	najdeni (TP)	napačno najdeni (FP)
brez pravil	472	923	100%	7330
s pravili	472	923	84.9%	304

Tabela 6.1: Uporaba metode na osnovi barv brez in z uporabo hevrističnih pravil na testni množici 472 barvnih slik, ki vsebujejo samo frontalne obraze.

obrazce na sivinskih slikah. S pomočjo te metode bom izločili kožne regije, ki ne vsebujejo obrazov in poskušal ohraniti kožne regije, ki vsebujejo obraz. Postopek iskanja frontalnih obrazov s pomočjo predlagane metode bom predstavili v nadaljevanju.

6.1 Postopek iskanja obraza

Sam postopek iskanja frontalnih obrazov na barvni sliki združuje postopek, s pomočjo katerega ohranimo kožno barvo, in postopek, s katerim iščemo na ohranjenih regijah frontalne obraze s pomočjo metode BDF. Postopek iskanja je sledeč:

Algoritem 1 Postopek iskanja frontalnega obraza na barvni sliki.

- vhodno sliko najprej ustrezno pomanjšamo na velikost 320×240 slikovnih elementov,
- na pomanjšani sliki ohranimo samo kožno barvo,
- ohranjeno barvo združimo v regije,
- s pomočjo preprostih pravil ohranimo samo regije ustreznih velikosti,

for all ohranjene regije **do**

for all različne velikosti **do**

for all različni nagibi **do**

 s pomočjo metode BDF išči frontalne obraze

end for

end for

end for

S pomočjo opisanega postopka lahko najdemo znotraj ohranjenih regij frontalne obraze različnih velikosti in nagibov, torej tudi obraze, ki imajo različen odklon od navpične lege.

V nadaljevanju bomo podrobno opisali posamezen korak predlaganega postopka iskanja obrazov.

Vhodno sliko najprej ustrezno zmanjšamo (slika 6.1a). Pri izbiri velikosti vhodne slike moramo biti pozorni na to, da izberemo dovolj veliko sliko, da lahko na njej najdemo tudi majhne obraze, ter da časovna kompleksnost metode ni prevelika. Pri izbrani velikosti 320×240 slikovnih elementov metoda na osnovi barve sliko obdela v manj kot eni sekundi. Optimizirana metoda BDF pa potrebuje za preiskovanje celotne slike v povprečju eno

sekundo. V večini primerov pa ne rabimo preiskovati celotne slike, saj je ohranjen samo majhen del slike, ki ustreza kožni barvi.

Na tako pomanjšani sliki nato izvedemo postopek iskanja kožne barve. V poglavju 4 smo predstavili tri metode, ki jih lahko uporabimo za ta postopek. V nadaljevanju bomo uporabili metodo na osnovi neparametričnega modela, ki se je izkazala za najuspešnejšo metodo. S pomočjo izbrane metode ohranimo samo kožno barvo, ki jo nato s pomočjo algoritma "rast regij" združimo v regije (slika 6.1b). Posamezno tako dobljeno regijo nato preverimo s pomočjo preprostih hevrističnih pravil (poglavje 4.2.1):

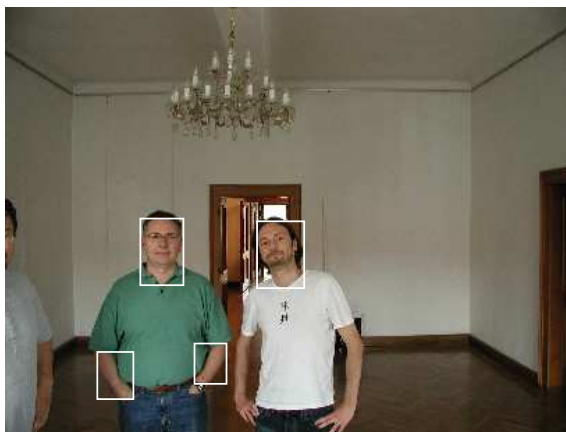


Slika 6.1: Postopek ohranjanja kožne barve: a) pomanjšana vhodna slika, b) ohranjena kožna barva združena v regije, c) ohranjene regije, ki zadoščajo pravilom.

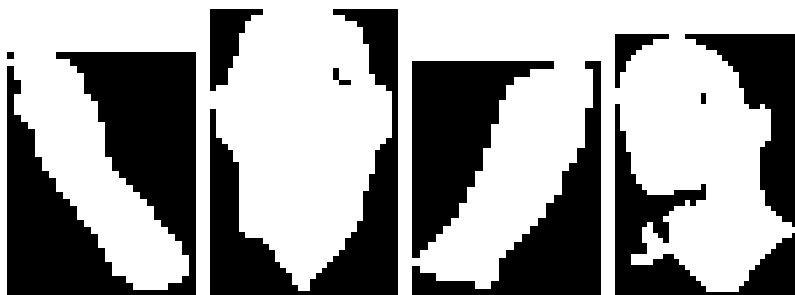
Pravila, ki smo jih uporabili v prejšnem preizkusu (tabela 6.1), smo nekoliko omilili, da ne izločijo preveč regij in s tem tudi kakšne regije, ki vsebuje obraz. Ohranjene regije (slika 6.2) nato uporabimo kot vhodne slike za metodo BDF. Na tako ohranjenih regijah s pomočjo metode BDF izvedemo postopek preiskovanja slike. Da še dodatno zmanjšamo preiskovalni čas, iščemo obraze samo na področju ohranjene kožne barve (slika 6.3).

Ker išče metoda BDF samo pokončne frontalne obraze standardne velikosti 16×16 slikovnih elementov, preiskujemo ohranjene regije pri različnih nagibih in različnih velikostih (slika 6.4). Na ta način lahko na sliki dobimo različno velike ter različno nagnjene obraze (slika 6.5).

Na posamezni regiji začnemo iskati obraze, ki ustrezajo velikostim širine same regije, pa do velikosti, ki ustreza polovici širine ohranjene regije. Na primer, če je širina ohranjene regije 60 slikovnih elementov, ohranjeno regijo zmanjšamo za faktor $16/60$ in jo po korakih velikosti 2 povečujemo do faktorja $16/30$, kjer lahko dobimo najmanjše obraze.

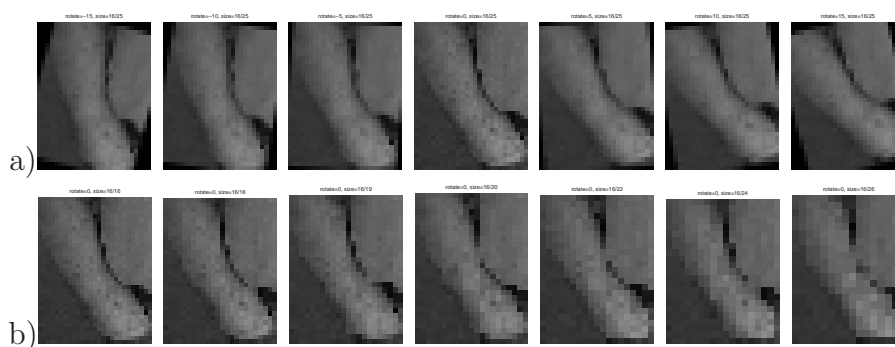


Slika 6.2: Vhodna slika z označenimi ohranjenimi kožnimi regijami.



Slika 6.3: Maske ohranjenih regij, ki predstavljajo ohranjeno kožno barvo.

Na takšen način nam ni potrebno preiskovati regije pri vseh različnih velikostih, ampak samo pri velikostih, ki ne odstopajo veliko od izbrane regije. Če obraza ne dobimo pri nobeni velikosti, regijo ustrezno nagnemo za korak petih stopinj. V našem primeru smo preiskovali obraze, ki so bili od navpičnega položaja odklonjeni za največ 15 stopinj (t.j. $\pm 5^\circ$, $\pm 10^\circ$ in $\pm 15^\circ$). Postopek preiskovanja regije zaključimo takoj, ko najdemo prvi obraz. Slika 6.6 prikazuje rezultat iskanja obrazov z ustrezno označenimi regijami, ki jih je ohranila metoda na osnovi barv (bel pravokotnik) in uspešno označenimi obrazi (rumen pravokotnik), ki jih je potrdila metoda BDF. Rdeči krogci označujejo ročno označene obraze, ki so nam služili za preizkus uspešnosti metode, ki je opisan v naslednjem poglavju.



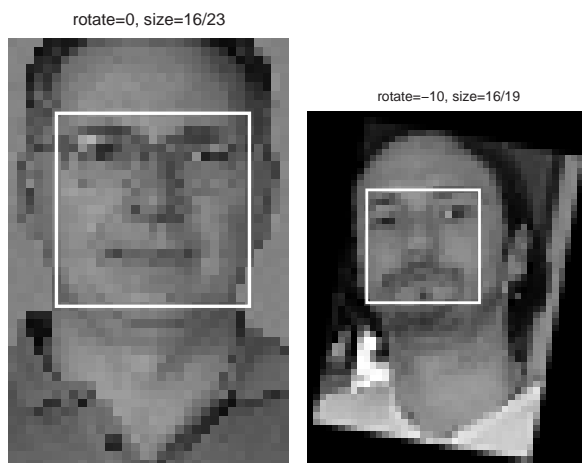
Slika 6.4: Ohranjena regija: a) pri različnih nagibih in b) različnih velikostih.

6.2 Preizkus metode in rezultati

Metodo smo preizkusili na testni množici barvnih slik, ki smo jih zajeli s pomočjo instalacije "15 sekund slave" [63]. Slike so bile posnete pod naravno osvetlitvijo. Med vsemi slikami smo izbrali samo tiste slike, na katerih so bili vsi obrazi v frontalnem položaju. Da smo lahko preizkusili uspešnost metode, smo posamezen obraz na sliki označili z rdečim krogcem v položaju nosu (slika 6.7).

Tabela 6.2 prikazuje rezultat iskanja frontalnih obrazov na slikah testne množice. Prva vrstica prikazuje rezultate po uporabi metode na osnovi barve. Vidimo lahko, da metoda na osnovi barv ohrani veliko regij, ki niso obrazi. Takšne regije so večinoma drugi deli telesa, ki so tudi kožne barve (roke, noge, vrat, itd.). Druga vrstica pa prikazuje končni rezultat metode po tem, ko izvedemo preiskovanje nad ohranjenimi regijami, ki smo jih ohranili s pomočjo metode na osnovi barv. Vidimo lahko, da metoda BDF močno zmanjša število napačno najdenih obrazov in še vedno ohrani visoko uspešnost. Če uporabimo samo metodo na osnovi barv in sicer z ostrejšimi hevrističnimi pravili, tudi zmanjšamo število napačno najdenih obrazov, ampak še vedno ne toliko, kot z našo kombinirano metodo (tabela 6.1).

Ker preiskuje metoda BDF samo znotraj ohranjenih regij lahko izračunamo tudi uspešnost same metode BDF. Metoda BDF med vsemi ohranjenimi regijami, med katerimi je 825 obrazov in 604 neobrazov, uspešno najde 765 obrazov in med neobrazi uspešno izloči 597 neobrazov. Torej je uspešnost (TP) metode BDF nad ohranjenimi regijami naše testne baze slik $765/825=92,7\%$ in uspešnost izločanja nepravilno ohranjenih regij (TN)



Slika 6.5: Primer dveh različno velikih in različno nagnjenih obrazov, ki jih je metoda uspešno našla.

metoda	št. slik	št. vseh obrazov	uspešnost (TP)	št. napačno najdenih (FP)
na osnovi barv	472	923	89.4%	604
kombinirana metoda	472	923	82.9%	7

Tabela 6.2: Uspešnost kombinirane metode na testni množici slik.

$597/604=98,8\%$.

Slika 6.8 prikazuje primere slik, na katerih je metoda BDF uspešno našla obraze. Metoda uspešno najde različno velike obraze, obraze pri različnih nagibih glave (odklon od navpičnega položaja), obraze na delno obrnjenih glavah (odklon od frontalnega položaja) in delno zakrite obraze. V vseh primerih lahko tudi vidimo, da je metoda uspešno izločila nepravilno ohranjene regije.

Slika 6.9 prikazuje najpogostejše napake pri iskanju obrazov z našo metodo. Pri neuspešno najdenih obrazih bomo tukaj obravnavali samo tiste, ki jih nismo našli s pomočjo metode BDF, čeprav jih je metoda na osnovi barve pravilno označila. Neuspešnost metod na osnovi barv pa podrobneje obravnava poglavje 4. Primer nepravilno izločene regije, ki ni vsebovala obraza, pa prikazuje slika 6.10.

Največkrat obraza nismo našli, ker je bil na sliki premajhen (slika 6.9a). Metoda BDF



Slika 6.6: Primer slike z uspešno najdenima obrazoma.

najde namreč samo obraze, ki so večji od 16×16 slikovnih elementov, kolikor je tudi standardna velikost vzorca. Vhodna slika, ki je v našem primeru velika 320×240 mora zato vsebovati dovolj velike obraze, da jih metoda še lahko najde. Druga najpogostejša napaka pa so obrazi, ki so preveč odstopali od frontalnega položaja (slika 6.9a). Ker je metoda BDF naučena, da išče samo frontalne obraze, je neuspešna pri iskanju obrazov, ki so preveč obrnjeni k profilnemu položaju. Veliko težav je imela metoda tudi pri osebah z očali, če je bil odsev očal prevelik (slika 6.9c). Takšen obraz je imel v položaju oči svetlejše slikovne elemente, kar ne ustreza vzorcu, ki ima v položaju oči temnejše slikovne elemente.

Težave pri iskanju so se pojavile tudi v primeru, če je bila označena regija, ki jo je označila metoda na osnovi barv, prevelika glede na velikost obraza znotraj te regije (slika 6.11). Ker je bil obraz znotraj te regije premajhen, ga nismo dobili, če smo preiskovali samo pri velikostih slike, ki so samo malo odstopale od velikosti označene regije. Težavo smo kasneje rešili tako, da smo takšne regije preiskovali tudi z manjšim vzorcem in uspešno našli obraze tudi v takšnih primerih.



Slika 6.7: Primer slike z ročno označenimi frontalnimi obrazi.



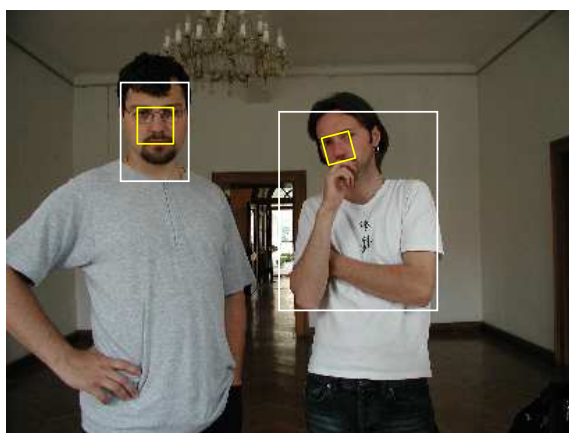
Slika 6.8: Primer slik z uspešno najdenimi obrazi: a) različno nagnjena glava, b) delno obrnjena glava proti profilnemu položaju, c) obrazi različnih velikosti, d) delno zakrit obraz.



Slika 6.9: Primer slik, na katerih metoda ni našla obraza: a) obraz je premajhen, b) obraz je preveč obrnjen, c) prevelik odsev ocal in d) kožna regija je napačno označena.



Slika 6.10: Primer slike, kjer je metoda nepravilno našla obraz.



Slika 6.11: Primer prevelike ohranjene kožne regije glede na velikost obraza.

Poglavje 7

Zaključek in nadaljnje delo

Na področju raziskav človeških obrazov, med katere spada tudi iskanje človeških obrazov na sliki, je bilo objavljenih že veliko del in predlaganih že veliko različnih metod. Zaradi tega v prvem poglavju naredim pregled vseh najaktualnejših metod in jih smiselno razporedim v štiri glavne skupine. Za vsako skupino opišem glavne predstavnike in na koncu podam rezultate za trenutno najuspešnejše metode, ki delujejo na osnovi videza.

Ker je bil moj cilj v tej magistrski nalogi poiskati obraze na barvnih slikah, sem kot prvo značilko, na osnovi katere sem poiskal obraze, uporabil kožno barvo. Ker je segmentacija barve odvisna od izbranega barvnega prostora, sem najprej opisal različne barvne prostore in njihove lastnosti. Izkazalo se je, da so za segmentacijo kožne barve najuspešnejši normalizirani kromatični barvni prostori, med katere spada tudi barvni prostor rg , ki sem ga sam uporabil. Za segmentacijo kožne barve sem preizkusil tri metode: metodo eksplicitno določenih mej, metodo na osnovi parametričnega modela in metodo na osnovi neparametričnega modela. Metodi na osnovi modela, ki za razliko od metode eksplicitno določenih mej uporabljata kromatični barvni prostor rg , sta se izkazali za boljši in bolj robustni pri segmentaciji kožne barve. Metodi sta zelo prilagodljivi in omogočata segmentacijo pri različnih osvetlitvah in različnih kožnih polteh. Ugotovil sem, da dosežemo najboljše rezultate, če se osvetlitev testnih slik ne razlikuje preveč od slik, na katerih je bil model naučen. Za najuspešnejšo metodo se je izkazala metoda na osnovi neparametričnega modela, ki jo lahko še dodatno izboljšamo z modelom nezaželenih barv. Na ta način lahko še dodatno zmanjšamo število napačno klasificiranih barv. Prednost

parametričnega modela glede na neparametrični model pa je ta, da potrebuje parametrični model manj učnih primerov ter je časovno in prostorsko manj zahteven.

Pri pregledu različnih metod za iskanje obrazov, so se metode na osnovi videza izkazale za najuspešnejše, zato sem realiziral metodo BDF, ki deluje na osnovi statistične metode razpoznavanja vzorcev. Metoda BDF uspešno najde frontalne človeške obraze na sivinskih slikah. Preizkusil sem jo na treh testnih množicah frontalnih obrazov in dosegel uspešnost 98.5%, z enim napačno najdenim obrazom. Metodo sem primerjal tudi s trenutno najuspešnejšo metodo na osnovi naivnega Bayesovega klasifikatorja in dosegel boljše rezultate, seveda na slikah frontalnih obrazov.

Metoda BDF se je pri iskanju frontalnih obrazov izkazala za zelo uspešno metodo. Ker sliko preiskujemo pri različnih velikostih in različnih nagibih, metoda uspešno najde tudi različno velike in različno nagnjene obraze. Uspešno najde tudi delno zakrite obraze, obraze na slikah slaše kvalitete in delno zasukane obraze. Ker pa je metoda naučena, da išče samo frontalne obraze, metoda ne najde obrazov s profila. Naslednja slabost metode je tudi njena časovna zahtevnost.

Z namenom, da bi zmanjšal časovno zahtevnost, sem v magistrskem delu izelal svojo metodo za iskanje frontalnih obrazov na barvnih slikah s kombinacijo dveh pristopov. S pomočjo prvega pristopa na osnovi barv sem ohranil samo kožno barvo in na tako segmentirani sliki se uporabil drugi pristop na osnovi videza oziroma metodo BDF. S kombinacijo obeh pristopov sem dosegel 24 krat hitrejše iskanje, kot če bi uporabil samo metodo BDF, ter izboljšal uspešnost iskanja frontalnih obrazov, kot če bi za to uporabil samo metodo na osnovi barv in hevrističnih pravil. Metodo sem tudi preizkusil na testni množici z 923 frontalnimi obrazi in dosegel uspešnost 82.9%, s sedmimi napačno najdenimi obrazi. Uspešnost je nekoliko manjša, kot če bi uporabili samo metodo BDF, to pa zato, ker je veliko obrazov napačno izločila že metoda na osnovi barv. Nekateri med ohranjenimi obrazi pa so bili premajhni ali pa niso bili pravilno izbrani zaradi slabe segmentacije kožne barve.

7.1 Nadaljnje delo

Metode na osnovi barev bi bilo potrebno preizkusiti na obsežnejši bazi obrazov različnih kožnih polti. Za takšen preizkus potrebujemo veliko obrazov različnih polti, da lahko zgradimo ustrezno veliki učno in testno množico.

Metodo BDF, ki trenutno uspešno išče samo frontalne obraze, bi bilo potrebno nadgraditi, da bi iskala tudi obraze s profila. Ker sem v magistrski nalogi izdelal metodo BDF s pomočjo jezika Matlab, bi bilo potrebno metodo izdelati v katerem od višjeprogramskih jezikov, kot sta C++, C# ali Pascal in s tem metodo pohitriti in optimizirati. S tako optimizirano metodo bi bili potem izvedljivi preizkusi nad obsežnejšimi podatkovnimi bazami slik. V magistrskem delu namreč nisem mogel preizkusiti uspešnosti metode BDF na testni množici 472 slik, ki sem jo uporabil za preizkus uspešnosti kombinirane metode. Preizkus na tej množici bi še dodatno pokazal, kakšen del doprinese oziroma kolikšen del uspešnosti izgubimo, če uporabimo novo kombinirano metodo. Iz istih razlogov tudi ne moremo določiti, kolikokrat hitrejša je nova kombinirana metoda od same metode BDF na testni množici slik. Vse to bi lahko izvedli samo s hitrejšo metodo BDF.

Moje nadaljnje raziskovanje bo obsegalo prepoznavo obrazov in prepoznavo mimike obraza. S pomočjo uporabljene metode bom na sliki poiskal obraz, ga izrezal in ga poskušal identificirati. S pomočjo prepoznavne mimike pa bom poskušal prepoznati razpoloženje osebe.

Zahvala

Zahvaljujem se mentorju prof. Fracu Solini za pomoč in podporo v času magistrskega študija ter pri pisanju magistrske naloge. Hvala prof. Igorju Kononenku in doc. Alešu Jakliču za popravke in predloge. Hvala tudi vsem, ki ste me s pravimi nasveti usmerili v raziskovanje te problematike, še posebej sodelavcema Petru Peeru in Katarini Mele.

Še posebej pa bi se rad zahvalil svoji zaročenki Tadeji, ki mi je tudi v najtežjih trenutkih stala ob strani in me podpirala.

Literatura

- [1] J. Ahlberg. *A system for face localization and facial feature extraction*. PhD thesis, Tech. Rep. LiTH-ISY-R-2172, Linköping University, 1999.
- [2] A. Albiol, L. Torres, and E. J. Delp. Optimum color spaces for skin detection. In *Proceedings of the International Conference on Image Processing*, volume 1, pages 122–124, 2001.
- [3] Y. Amit, D. Geman, and B. Jedynek. Efficient focusing and face detection. In V. Bruce F. Fogelman-Soulie H. Wechsler, P.J. Phillips and T.S. Huang, editors, *Face Recognition: From Theory to Applications*, volume 163, pages 124–156. 1998.
- [4] Marijke F. Augusteijn and Tammy L. Skufca. Identification of human faces through texture-based feature recognition and neural network technology. In *Proceedings of 1993 IEEE International Conference on Neural Networks*, volume I, pages 392–398, San Francisco, California, March-April 1993. IEEE/INNS.
- [5] J. Brand and J. Mason. A comparative assessment of three approaches to pixellevel human skin-detection. In *Proc. of the International Conference on Pattern Recognition*, pages 1056–1059, 2000.
- [6] Leo Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. *Classification and Regression Trees*. Statistics/Probability Series. Wadsworth Publishing Company, Belmont, California, U.S.A., 1984.

- [7] D. Brown, I. Craw, and J. Lewthwaite. A som based approach to skin detection with application in real time systems. In *Proc. of the British Machine Vision Conference*, 2001.
- [8] Gilles Burel and Dominique Carel. Detection and localization of faces on digital images. *Pattern Recogn. Lett.*, 15(10):963–967, 1994.
- [9] MC. Burl, T.K. Leung, and P. Perona. Face Localization via Shape Statistics. In *Proc. First Int’l Workshop Automatic Face and Gesture Recognition*, pages 154–159, 1995.
- [10] D. Chai and A. Bouzerdoun. A bayesian approach to skin color classification in ycbcr color space. In *Proceedings IEEE Region Ten Conference (TENCON’2000)*, volume 2, pages 421–424, 2000.
- [11] Q. Chen, H. Wu, and M. Yachida. Face detection by fuzzy matching. In *Proc. Fifth IEEE Int’l Conf. Computer Vision*, pages 591–596, 1995.
- [12] D. Chetverikov and A. Lerch. Multiresolution face detection. *Theoretical Foundations of Computer Vision*, 69:131–140, 1993.
- [13] I. Craw, H. Elis, and J. R. Lishman. Automatic extraction of face-features. *Pattern Recogn. Lett.*, 5(2):183–187, 1987.
- [14] Ian Craw, David Tock, and Alan Bennett. Finding face features. In *Proceedings of the Second European Conference on Computer Vision*, pages 92–96. Springer-Verlag, 1992.
- [15] J. Crowley and F. Berard. Multi-modal tracking of faces for video communication. In *Proc. IEEE Conf Computer Vision and Pattern Recognition*, pages 640–645, 1997.
- [16] J. L. Crowley, J. M. Bedrune, M. Bekker, and M. Schneider. Integration and control of reactive visual processes. *Lecture Notes in Computer Science*, 800:47–60, 1994.
- [17] C.Wang and M.Brandstein. Multi-source face tracking with audio and visual data. In *IEEE MMSP*, pages 169–174, 1999.

- [18] Y. Dai and Y. Nakano. Face-texture model based on sgld and its application in face detection in a color scene. *Pattern Recognition*, 29(6):1007–1017, 1996.
- [19] M. Fleck, D. A. Forsyth, and C. Bregler. Finding naked people. In *Proc. of the ECCV*, volume 2, pages 592–602, 1996.
- [20] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 23(6):643–660, 2001.
- [21] G. Gomez. On selecting colour components for skin detection. In *Proc. of the ICPR*, volume 2, pages 961–964, 2000.
- [22] G. Gomez and E. Morales. Automatic feature construction and a simple rule induction algorithm for skin detection. In *Proc. of the ICML Workshop on Machine Learning in Computer Vision*, pages 31–38, 2002.
- [23] H. P. Graf, T. Chen, F. Petajan, and E. Cosatto. Locating Faces and Facial Parts. In *Proc. First Int’l Workshop Automatic Face and Gesture Recognition*, pages 41–46, 1995.
- [24] C.-C. Han, H.-Y.M. Liao, K.-C. Yu, and L.-H. Chen. Fast Face Detection via Morphology-Based Pre-Processing. In *Proc. Ninth Int’l Conf. Image Analysis and Processing*, pages 469–476, 1998.
- [25] R.-L. Hsu, M. Abdel-Mottaleb, and A. K. Jain. Face detection in color images. In *IEEE Trans. Pattern Analysis and Machine Intelligence* 24, volume 5, pages 696–706, 2002.
- [26] T.S. Jebara and A.P. Pentland. Parameterized structure from motion for 3d adaptive feedback tracking of faces. In *CVPR97*, pages 144–150, 1997.
- [27] I. T. Jolliffe. *Principal Component Analysis*. Springer, 1986.
- [28] M. J. Jones and J. M. Rehg. Statistical color models with application to skin detection. In *Proceedings of the IEEE Computer Science Conference on Computer Vision*

- and Pattern Recognition (CVPR-99)*, pages 274–280, Los Alamitos, June 23–25 1999. IEEE.
- [29] M.J. Jones and J.M. Rehg. Statistical color models with application to skin detection. In *CVPR99*, volume 1, pages 274–280, 1999.
- [30] L. Jordao, M. Perrone, J. Costeira, and J. Santos-Victor. Active face and feature tracking. In *Proceedings of the 10th International Conference on Image Analysis and Processing*, pages 572–577, 1999.
- [31] T. Kanade. *Picture Processing by Computer Complex and Recognition of Human Faces*. PhD thesis, Kyoto Univ., 1973.
- [32] D. G. Kendall. Shape manifolds, procrustean metrics, and complex projective spaces. *Bull. London Math. Soc.*, 16:81–121, 1984.
- [33] M. Kirby and L. Sirovich. Application of the karhunen-loeve procedure for the characterization of human faces. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(1):103–108, 1990.
- [34] T. Kohonen. *Self-organization and associative memory: 3rd edition*. Springer-Verlag New York, Inc., 1989.
- [35] C. Kotropoulos and I. Pitas. Rule-based face detection in frontal views. In *Proceedings of IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP-97)*, volume 4, pages 2537–2540, Munich, Germany, April 21–24 1997. IEEE.
- [36] A. Lanitis, C.J. Taylor, and T.F. Cootes. Automatic face identification system using flexible appearance models. *IVC*, 13(5):393–401, June 1995.
- [37] J. Y. Lee and S. I. Yoo. An elliptical boundary model for skincolor detection. In *Proc. of the 2002 International Conference on Imaging Science, Systems, and Technology*, 2002.
- [38] T.K. Leung, M. C. Burl, and P. Perona. Finding faces in cluttered scenes using random labeled graph matching. In *Fifth Intl. Conf. on Comp. Vision*, pages 637–644, June 1995.

- [39] T.K. Leung, M.C. Burl, and P. Perona. Probabilistic Affine Invariants for Recognition. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 678–684, 1998.
- [40] C. Liu. A bayesian discriminating features method for face detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(6):725–740, 2003.
- [41] K. V. Mardia and I. L. Dryden. Shape distributions for landmark data. *Adv. Appl. Probab.*, 21:742–755, 1989.
- [42] F. Marques and V. Vilaplana. A morphological approach for segmentation and tracking of human faces. In *International Conference on Pattern Recognition (ICPR'00)*, volume 1, pages 5064–5068, 2000.
- [43] S. McKenna, S. Gong, and Y. Raja. Modelling facial colour and identity with gaussian mixtures. In *Pattern Recognition*, volume 12, pages 1883–1892, 1998.
- [44] S.J. McKenna, Y. Raja, and S. Gong. Tracking colour objects using adaptive mixture models. *IVC*, 17(3/4):225–231, March 1999.
- [45] B. Menser and M. Wien. Segmentation and tracking of facial regions in color image sequences. In *Proc. SPIE Visual Communications and Image Processing*, pages 731–740, 2000.
- [46] Baback Moghaddam and Alex Pentland. Probabilistic visual learning for object representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):696–710, 1997.
- [47] N. Oliver, A. Pentland, and F. Berard. Lafter: Lips and face real time tracker. In *Proc. Computer Vision and Pattern Recognition*, pages 123–129, 1997.
- [48] E. Osuna, R. Freund, and F. Girosi. Training support vector machines: An application to face detection. In *Proc. Computer Vision and Pattern Recognition '97*, pages 130–136, 1997.
- [49] P.J. Phillips, H. Wechsler, J. Huang, and P. Rauss. The feret database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing*, 16:295–306, 1998.

- [50] S. L. Phung, A. Bouzerdoum, and D. Andchai. A novel skin color model in ycbcr color space and its application to human face detection. In *IEEE International Conference on Image Processing(ICIP'2002)*, pages 289–292, 2002.
- [51] S. Pigeon and L. Vandendrope. The M2VTS Multinodal Face Database. In *Proc. First Int'l Conf. Audio- and Video-Based Biometric Person Authentication*, 1997.
- [52] C. A. Poynton. Frequently asked questions about colour. In *ftp://www.inforamp.net/pub/users/poynton/doc/colour/ColorFAQ.ps.gz*, 1997.
- [53] H. A. Rowley, S. Baluja, and T. Kanade. Rotation invariant neural network-based face detection. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, page 38. IEEE Computer Society, 1998.
- [54] Henry A. Rowley, Shumeet Baluja, and Takeo Kanade. Human face detection in visual scenes. In David S. Touretzky, Michael C. Mozer, and Michael E. Hasselmo, editors, *Advances in Neural Information Processing Systems*, volume 8, pages 875–881. The MIT Press, 1996.
- [55] Henry A. Rowley, Shumeet Baluja, and Takeo Kanade. Neural network-based face detection. In *Proceedings of the 1996 Conference on Computer Vision and Pattern Recognition (CVPR '96)*, page 203. IEEE Computer Society, 1996.
- [56] E. Saber and A. Tekalp. Frontal-view face detection and facial feature extraction using color, shape and symmetry based cost functions. *Pattern Recognition Letters*, 17(8):669–680, 1998.
- [57] T. Sakai, M. Nagao, and S. Fujibayashi. Line extraction and pattern detection in a photograph. *PR*, 1(3):233–236, March 1969.
- [58] Ferdinando Samaria. *Face Recognition Using Hidden Markov Models*. PhD thesis, Engineering Department, Cambridge University, oct 1994.
- [59] H. Schneiderman and T. Kanade. Probabilistic modeling of local appearance and spatial relationships for object recognition. In *Proc IEEE Conf. Computer Vision and Pattern Recognition*, pages 45–51, 1998.

- [60] R. Schumeyer and K. Barner. A color-based classifier for region identification in video. In *Visual Communications and Image Processing 1998, SPIE*, volume 3309, pages 189–200, 1998.
- [61] S. A. Sirohey. Human face segmentation and identification. Technical Report CS-TR-3176, University of Maryland Computer Vision Laboratory, 1993.
- [62] K. Sobottka and I. Pitas. Face localization and feature extraction based on shape and color information. In *Proc. IEEE Int'l Conf. Image Processing*, pages 483–486, 1996.
- [63] F. Solina, P. Peer, B. Batagelj, S. Juvan, and J. Kovac. Color-Based Face Detection in the "15 Seconds of Fame" Art Installation. In *International Conference on Computer Vision / Computer Graphics Collaboration for Model-based Imaging, Rendering, image Analysis and Graphical special Effects MIRAGE'03*, pages 38–47, 2003.
- [64] Franc Solina. 15 seconds of fame. *Leonardo (Oxf.)*, 37(2):105–110,125, 2004.
- [65] K.-K. Sung. *Learning and Example Selection for Object and Pattern Detection*. PhD thesis, Massachusetts Inst. of Technology, 1996.
- [66] Kah K Sung and Tomaso Poggio. Example based learning for view-based human face detection. Technical report, 1994.
- [67] J. C. Terrillon, M. David, and S. Akamatsu. Automatic detection of human faces in natural scene images by use of a skin color model and of invariant moments. In *Proceedings of the 3rd. International Conference on Face and Gesture Recognition*, pages 112–117. IEEE Computer Society, 1998.
- [68] J.C. Terrillon, M. David, and S. Akamatsu. Detection of human faces in complex scene images by use of a skin color model and of invariant fourier-mellin moments. In *ICPR98*, pages 1350–1355, 1998.

- [69] J.C. Terrillon, M.N. Shirazi, H. Fukamachi, and S. Akamatsu. Comparative performance of different skin chrominance models and chrominance spaces for the detection of human faces in color images. In *AFGR00*, pages 54–61, 2000.
- [70] M. Turk and A.P. Pentland. Eigenfaces for recognition. *CogNeuro*, 3(1):71–96, 1991.
- [71] R. Vaillant, C. Monrocq, and Y. Le Cun. An Original Approach for the Localisation of Objects in Images. In *IEEE Proc. Vision, Image and Signal Processing*, volume 141, pages 245–250, 1994.
- [72] T. Vene. Analiza učinkovitosti in izboljšava algoritma za iskanje obrazov na slikah. Diplomsko delo, Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, april 2003.
- [73] T. Wark and S. Sridharan. A syntactic approach to automatic lip feature extraction for speaker identification. In *ICASSP*, pages 3693–3696, 1998.
- [74] Haiyuan Wu, Qian Chen, and Masahiko Yachida. Face detection from color images using a fuzzy pattern matching method. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21(6):557–563, 1999.
- [75] Haiyuan Wu, Taro Yokoyama, Dadet Pramadihanto, and Masahiko Yachida. Face and facial feature extraction from color image. In *Proceedings of the 2nd International Conference on Automatic Face and Gesture Recognition (FG '96)*, page 345. IEEE Computer Society, 1996.
- [76] G. Yang and T. S. Huang. Human face detection in complex background. *Pattern Recognition*, 27(1):53–63, 1994.
- [77] J. Yang, W. Lu, and A. Waibel. Skin-color modeling and adaptation. In *Proceedings of ACCV 1998*, pages 687–694, 1998.
- [78] Jie Yang and unknown. A real-time face tracker. In *Proceedings of WACV'96*, 1996.
- [79] M. Yang and N. Ahuja. Gaussian mixture model for human skin color and its application in image and video databases. In *Proc. SPIE: Storage and Retrieval for Image and Video Databases VII*, volume 3656, pages 458–466, 1999.

- [80] M.-H. Yang and N. Ahuja. Detecting human faces in color images. In *International Conference on Image Processing (ICIP)*, volume 1, pages 127–130, 1998.
- [81] Ming-Hsuan Yang, David J. Kriegman, and Narendra Ahuja. Detecting faces in images: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(1):34–58, 2002.
- [82] K.C. Yew and R. Cipolla. A Probabilistic Framework for Perceptual Grouping of Features for Human Face Detection. In *Proc. Second Int’l Conf. Automatic Face and Gesture Recognition*, pages 16–21, 1996.
- [83] K.C. Yew and R. Cipolla. Feature-based human face detection. *Image and Vision Computing*, 15(9):713–735, 1997.
- [84] B. Zarit, B. J. Super, and F. Quek. Comparison of five color models in skin pixel classification. In *ICCV’99 Int’l Workshop on recognition, analysis and tracking of faces and gestures in Real-Time systems*, pages 58–63, 1999.

Izjava

Izjavljam, da sem magistrsko delo izdelal samostojno pod vodstvom mentorja prof. dr. Franca Soline. Izkazano pomoč drugih sodelavcev sem v celoti navedel v zahvali.